

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 17:47:33 ; Search time 37 Seconds
(without alignments)
3246.634 Million cell updates/sec

Title: US-09-924-654-4

Perfect score: 3209
Sequence: 1 MSYGEFGKFLGPREVTSE.....RQRMGSLGSHNRVCNTNH 583

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3209	100.0	583	4 Q92802	Q92802 homo sapien
2	1975	61.5	575	11 Q912G4	Q912G4 mus musculu
3	509	15.9	659	4 Q9NWK2	Q9NWK2 homo sapien
4	469	14.6	1399	4 Q9P2D4	Q9P2D4 homo sapien
5	449.5	14.0	87	4 Q43712	Q43712 homo sapien
6	427	13.3	238	11 Q912G6	Q912G6 mus musculu
7	334	10.4	669	5 Q9X211	Q9X211 drosophila
8	278.5	8.7	145	11 Q912G7	Q912G7 mus musculu
9	272.5	8.5	173	4 Q8WTU5	Q8WTU5 homo sapien
10	247	7.7	137	4 Q9Y273	Q9Y273 homo sapien
11	242.5	7.6	134	11 Q8VE50	Q8VE50 mus musculu
12	195.5	6.1	232	16 Q25201	Q25201 helicobacte
13	164.5	5.1	162	5 Q19804	Q19804 caenorhabdi
14	161	5.0	149	16 Q9KRM6	Q9KRM6 vibrio chol
15	147	4.6	400	11 Q91V42	Q91V42 mus musculu
16	147	4.6	420	11 Q923F3	Q923F3 mus musculu

17	145.5	4.5	644	10 Q9FM17	Q9FM17 arabidopsis
18	143	4.5	3978	5 Q97236	Q97236 plasmodium
19	141	4.4	1078	5 Q96371	Q96371 plasmodium
20	136.5	4.3	1551	10 Q9SYPO	Q9SYPO arabidopsis
21	136	4.2	649	4 Q9H0T2	Q9H0T2 homo sapien
22	136	4.2	944	5 Q17411	Q17411 caenorhabdi
23	135.5	4.2	420	13 Q57389	Q57389 gallus gall
24	135	4.2	416	5 Q9U006	Q9U006 plasmodium
25	134.5	4.2	2269	5 Q77360	Q77360 plasmodium
26	134	4.2	1189	5 Q967Q7	Q967Q7 plasmodium
27	134	4.2	1336	5 Q9GTX2	Q9GTX2 plasmodium
28	134	4.2	1271	5 Q25860	Q25860 plasmodium
29	134	4.2	1695	5 Q9BKL2	Q9BKL2 hydra atten
30	133	4.1	420	11 Q64575	Q64575 rattus norv
31	132	4.1	2747	5 Q9B3X9	Q9B3X9 plasmodium
32	131	4.1	416	5 Q9TVT7	Q9TVT7 plasmodium
33	131	4.1	1933	13 Q90337	Q90337 cyprinus ca
34	131	4.1	3620	5 Q968T6	Q968T6 plasmodium
35	130.5	4.1	2485	5 Q96134	Q96134 plasmodium
36	130	4.1	416	5 Q9U0R0	Q9U0R0 plasmodium
37	130	4.1	416	5 Q9U0Q5	Q9U0Q5 plasmodium
38	130	4.1	416	5 Q9TVJ5	Q9TVJ5 plasmodium
39	130	4.1	1946	5 Q97291	Q97291 plasmodium
40	129	4.0	416	5 Q9U0Q7	Q9U0Q7 plasmodium
41	129	4.0	416	5 Q9TVJ2	Q9TVJ2 plasmodium
42	128	4.0	416	5 Q9TW66	Q9TW66 plasmodium
43	128	4.0	2500	5 Q96223	Q96223 plasmodium
44	128	4.0	2924	5 Q25733	Q25733 plasmodium
45	128	4.0	4226	5 Q9N9H5	Q9N9H5 plasmodium

ALIGNMENTS

RESULT 1

ID Q92802 PRELIMINARY: PRT: 583 AA.

AC Q92802; 01-FEB-1997 (TREMBLrel. 02, Created)

DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Hypothetical 67.5 kDa protein.

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96411650; PubMed=8812419;

RA Couch F.J., Rommens J.M., Neuhausen S.L., Belanger C., Dumont M.,

RA Kenneth A., Bell R., Berry S., Bogden R., Cannon-Albright L.,

RA Farid L., Frye C., Hattler T., Janacki T., Jiang P., Kehrer R.,

RA LeBlanc J.F., McArthur-Morrison J., McSweney D., Miki Y., Peng Y.,

RA Samsen C., Schroeder M., Snyder S.C., Stringfellow M., Stroup C.,

RA Swedlund B., Swensen J., Teng D., Thakur S., Tran T., Tranchesi M.,

RA Weller-Feldhaus J., Wong A.K.C., Shizuya H., Labrie F., Skolnick M.H.,

RA Goldgar D.E., Kamb A., Weber B.L., Tavtigian S.V., Simard J.,

RT "Generation of an integrated transcription map of the BRCA2 region on

chromosome 13q12-q13."

RL Genomics 36:86-99(1996).

RN [2]

RP SEQUENCE FROM N.A.

RA Rhodes S.;

RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RA EMBL: U50532; AAC50875.1; -

DR EMBL: AL049783; CAB42441.1; -

DR EMBL: BC010643; AAH10643.1; -

KW Hypothetical protein.

```
SQ SEQUENCE 583 AA; 67459 MW; 318DC4D81CD0FF2A CRC64;
Query Match 100.0%; Score 3209; DB 4; Length 583;
Best Local Similarity 100.0%; Pred. No. 7,7e-212;
Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSYGEIEGKFLGPREVTSERCKKLKSTTESYVFHNSNADFHRIQKTCNDWVPVTII 60
DB 1 MSYGEIEGKFLGPREVTSERCKKLKSTTESYVFHNSNADFHRIQKTCNDWVPVTII 60
QY 61 DVRGHSYLQENKIKTTDLRPLHDEMPGNRPDVIESIDSVLQEARPLVSADDEIYSTS 120
DB 61 DVRGHSYLQENKIKTTDLRPLHDEMPGNRPDVIESIDSVLQEARPLVSADDEIYSTS 120
QY 121 KAFITGPIYKPEKKRNRGRNEAHVNLGINDRGGQKQKFNSEKSEIDNLFQFYKIE 180
DB 121 KAFITGPIYKPEKKRNRGRNEAHVNLGINDRGGQKQKFNSEKSEIDNLFQFYKIE 180
QY 181 ELEKEKDFGENSEKSESPSQOFVFPFYEGHNNGLLKPDDEKKDLSNKAMPSCDYQOQNLG 240
DB 181 ELEKEKDFGENSEKSESPSQOFVFPFYEGHNNGLLKPDDEKKDLSNKAMPSCDYQOQNLG 240
QY 241 NEPKYPCNGOVIPFTCDTSTSPRPQSVYFIVPYGPPPLSLNYHLNLTQRFSGPPNP 300
DB 241 NEPKYPCNGOVIPFTCDTSTSPRPQSVYFIVPYGPPPLSLNYHLNLTQRFSGPPNP 300
QY 301 PSNIFQAODDSQIQNGYVYNNCHVNNWCMTFDQNNETDCSENSSVHPSCGSMQDRY 360
DB 301 PSNIFQAODDSQIQNGYVYNNCHVNNWCMTFDQNNETDCSENSSVHPSCGSMQDRY 360
QY 361 VSNCFCEVREKWCMDKCHNGTDRFVNOQFOBEKLNKLOKLLILLRGLPGSGKTTLSR 420
DB 361 VSNCFCEVREKWCMDKCHNGTDRFVNOQFOBEKLNKLOKLLILLRGLPGSGKTTLSR 420
QY 421 ILLGNRDRGIVFSTDDYFHHQDGYRYNNQLGDAHDNQNRAKQAIQDGRSPVIIDNTNI 480
DB 421 ILLGNRDRGIVFSTDDYFHHQDGYRYNNQLGDAHDNQNRAKQAIQDGRSPVIIDNTNI 480
QY 481 QAWEMKPYVEVAIGKYRVEFHEPETWKKFDPBELEKRNKHGVSRRKKTQAQMLDRYEYQMS 540
DB 481 QAWEMKPYVEVAIGKYRVEFHEPETWKKFDPBELEKRNKHGVSRRKKTQAQMLDRYEYQMS 540
QY 541 ISIVMNSVEPSHKSTQRPPTQPGQRWGGSLGSHNRVCVTNNH 583
DB 541 ISIVMNSVEPSHKSTQRPPTQPGQRWGGSLGSHNRVCVTNNH 583

RESULT 2
QY12G4 ID Q91ZG4 PRELIMINARY; PRT; 575 AA.
AC Q91ZG4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE DM417G6.4 (novel protein).
GN DM417G6.4
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wall M.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL358892; CAC42205.1;
SQ SEQUENCE 575 AA; 66250 MW; C41FAA8CC3EE4408 CRC64;

Query Match 61.5%; Score 1975; DB 11; Length 575;
Best Local Similarity 65.2%; Pred. No. 2,7e-127;
Matches 380; Conservative 70; Mismatches 125; Indels 8; Gaps 7;

QY 1 MSYGEIEGKFLGPREVTSERCKKLKSTTESYVFHNSNADFHRIQKTCNDWVPVTII 60
DB 1 MSYGEIEGKFLGPREVTSERCKKLKSTTESYVFHNSNADFHRIQKTCNDWVPVTII 60
```

```
DB 1 MPYSEVEAKFLGPKGEQTRPCYKLLKSAADDGVSPLRGDPDIHRIQKPRNNRVAVATI 60
QY 61 DVRGHSYLQENKIKTTDLRPLHDEMPGNRPDVIESIDSVLQEARPLVSADDEIYSTS 120
DB 61 DVRGHSYLQENKIKTTDLRPLHDEMPGNRPDVIESIDSVLQEARPLVSADDEIYSTS 120
QY 121 KAFITGPIYKPEKKRNRGRNEAHVNLGINDRGGQKQKFNSEKSEIDNLFQFYKIE 180
DB 121 KAFITGPIYKPEKKRNRGRNEAHVNLGINDRGGQKQKFNSEKSEIDNLFQFYKIE 180
QY 181 ELEKEKDFGENSEKSESPSQOFVFPFYEGHNNGLLKPDDEKKDLSNKAMPSCDYQOQNLG 240
DB 181 ELEKEKDFGENSEKSESPSQOFVFPFYEGHNNGLLKPDDEKKDLSNKAMPSCDYQOQNLG 240
QY 241 NEPKYPCNGOVIPFTCDTSTSPRPQSVYFIVPYGPPPLSLNYHLNLTQRFSGPPNP 300
DB 241 NEPKYPCNGOVIPFTCDTSTSPRPQSVYFIVPYGPPPLSLNYHLNLTQRFSGPPNP 300
QY 301 PSNIFQAODDSQIQNGYVYNNCHVNNWCMTFDQNNETDCSENSSVHPSCGSMQDRY 360
DB 301 PSNIFQAODDSQIQNGYVYNNCHVNNWCMTFDQNNETDCSENSSVHPSCGSMQDRY 360
QY 361 VSNCFCEVREKWCMDKCHNGTDRFVNOQFOBEKLNKLOKLLILLRGLPGSGKTTLSR 420
DB 361 VSNCFCEVREKWCMDKCHNGTDRFVNOQFOBEKLNKLOKLLILLRGLPGSGKTTLSR 420
QY 421 ILLGNRDRGIVFSTDDYFHHQDGYRYNNQLGDAHDNQNRAKQAIQDGRSPVIIDNTNI 480
DB 421 ILLGNRDRGIVFSTDDYFHHQDGYRYNNQLGDAHDNQNRAKQAIQDGRSPVIIDNTNI 480
QY 481 QAWEMKPYVEVAIGKYRVEFHEPETWKKFDPBELEKRNKHGVSRRKKTQAQMLDRYEYQMS 540
DB 481 QAWEMKPYVEVAIGKYRVEFHEPETWKKFDPBELEKRNKHGVSRRKKTQAQMLDRYEYQMS 540
QY 541 ISIVMNSVEPSHKSTQRPPTQPGQRWGGSLGSHNRVCVTNNH 583
DB 541 ISIVMNSVEPSHKSTQRPPTQPGQRWGGSLGSHNRVCVTNNH 583

RESULT 3
QY9NVK2 ID Q9NVK2 PRELIMINARY; PRT; 659 AA.
AC Q9NVK2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CDNA FLJ10680 fis, clone NT2RP2006573, weakly similar to 2',3'-cyclic
DE nucleotide 3'-phosphodiesterase (EC 3.1.4.37) (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA Masubo Y., Kanehori K.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001542; BAA91748.1;
DR InterPro; IPR001899; Gram_pos_anchor.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
FT NON_TER 659
SQ SEQUENCE 659 AA; 73887 MW; 9974015C7D6EAC1F CRC64;

Query Match 15.9%; Score 509; DB 4; Length 659;
Best Local Similarity 27.9%; Pred. No. 9,2e-27;
Matches 165; Conservative 88; Mismatches 157; Indels 182; Gaps 28;

QY 45 RIQKTCNDWVPVTIIIVRGHSYLQENKIKTTD---LHRPLHDEMPGNRPD-----VIE 95
DB 45 RIQKTCNDWVPVTIIIVRGHSYLQENKIKTTD---LHRPLHDEMPGNRPD-----VIE 95
```

```
Db 93 KIEESSOSFV-----ASENOYGAESKIMEKREPEESDSKMSDFLQMLTE 140
Qy 96 SIDSOVLDEAPPLVSA--DEIYS-----TSKAFIPPIYKPKKKRNEGREA 143
Db 141 DIDS-LIQNAFEKLNSSPDQVSEFLPSQDVNSFNSSEFINP-----SSNMT 188
Qy 144 HV-----LNCIN-DRGQ-----KEKQKFNSEKSEI-DNELF--QFYKEIETL--- 182
Db 189 PIFSTQNMNLNGENLSSSTLSLNPFLPSHVSVLNESKCFIKNTALLESNPEDSILSS 248
Qy 183 -----EKEKGFENSKCE-----SEPSOEQVFPFEGHNNGLL 215
Db 249 LNVASDSIAGCSLNGKQKELLESECEVGAQFSEAPYDLASBPQACLNLP-----GLD 301
Qy 216 KP---DEEKKDLNKKAMPSSHCDOQONIGNEPDKYPCNGQVIPT-----FCDTSFTSR 265
Db 302 LPTGGDQKSTRVSDVFLPS-----EGFNFKPKHKP---ELPTGKGDVSYCPV----- 346
Qy 266 PEMQSYPIPIVYGPPLPLNHLNIGRSGPPNPSPNIFQADDSQIONGY--VNNCH 323
Db 347 ---LAPFLPLPPPPPPPMW-----PMIPAFDLFG-----NHGFVAPVVTGA 387
Qy 324 VMNMGTFQNNNEYTPDCSENSSVHPSGNGCSMDRYVNGCEVRECMKDHCKMKG 383
Db 388 AHMRSV-----NYTPPPSVISHTSPT-----KYMRN-----KDG 416
Qy 384 TDFRVNQ---FO--EEKLNKLQKLLILRLGSLPGSGTTLIRLLGQNRDGIYSTDYF 438
Db 417 TSAIYQVETFPVQVYARKTSYGLVLYLRGLPGSGKSLARLQEDNDSVILSTDYF 476
Qy 439 HHODGIRYVNOIGDAHNMNONAKAIDOGSPVIIDNTINQAMKRYVEVAIKGR 498
Db 477 YINGQYQFVKKYLGEAHENNONAKAEFEKKSPIIIDNTINQAMKRYVLSQHKYK 536
Qy 499 VERHEBETWKKDPPELEKRNKGVSRRKKIAQMLDREYOMSISIVANSSEP 550
Db 537 VLFREPDYMKRKPFLARNIHGVSKKITRLEHYQRFVSVIITSSSVP 588
```

RESULT 4

```
Q9P2D4 PRELIMINARY: PRT: 1399 AA.
AC Q9P2D4: 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE KIAA1413 protein (Fragment).
GN KIAA1413.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 7:65-73(2000).
DR EMBL: AB037834; BAA92651.1;
DR InterPro: IPR002625; Smr/MutS2.
DR Pfam: PF01713; Smr. 1.
DR SMART: SM00463; Smr. 1.
FT NON_TER 1
SQ SEQUENCE 1399 AA: 158124 MW: 10ADC0B597C22BE0 CRC64;
```

Query Match 14.6%; Score 469; DB 4; Length 1399;
Best Local Similarity 47.6%; Pred. No. 1.3e-23;
Matches 88; Conservative 40; Mismatches 47; Indels 10; Gaps 2;
Qy 371 RCKWCHMCKHNGCTDFRVNO-----OFEBKLNKLQKLLILRLGSLPGSGTTLIRLLIGQ 425

```
Db 55 KYMRN-----KGTSAIYQVETFPVQVYARKTSYGLVLYLRGLPGSGKSLARLQED 109
Qy 426 NEDGIVFSTDYFHHODGIRYVNOIGDAHNMNONAKAIDOGSPVIIDNTINQAMK 485
Db 110 NPSGVLSTDYFYINGQYQFVKKYLGEAHENNONAKAEFEKKSPIIIDNTINQAMK 169
Qy 486 KPYVEAIGKGYRVFHEBETWKKDPPELEKRNKGVSRRKKIAQMLDREYOMSISIVM 545
Db 170 KPYVALSQKHKKYLVLFREPDYMKRKPFLARNIHGVSKKITRLEHYQRFVSVIIM 229
Qy 546 NSVEP 550
Db 230 SSSVP 234
```

RESULT 5

```
043712 PRELIMINARY: PRT: 87 AA.
AC 043712: 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE XS7 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mcmurray A., Hunt A.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z75887; CAB00067.1;
FT NON_TER 87
SQ SEQUENCE 87 AA: 10074 MW: 2E350DF672BA3CBA CRC64;
```

Query Match 14.0%; Score 449.5; DB 4; Length 87;
Best Local Similarity 82.1%; Pred. No. 8.5e-24;
Matches 87; Conservative 0; Mismatches 0; Indels 19; Gaps 1;

```
Qy 1 MSYGEIEGKFLGPREEVTSPPCKKIKSTSEYVFHNSNADFHRIQETGNDWVPVIT 60
Db 1 MSYGEIEGKFLGPRE-----HNHNSNADFHRIQETGNDWVPVIT 41
Qy 61 DVGHSYIQENKIKTTDLRPLHDEMPGRPDVIESIDSOVLQEAR 106
Db 42 DVGHSYIQENKIKTTDLRPLHDEMPGRPDVIESIDSOVLQEAR 87
```

RESULT 6

```
Q91ZG6 PRELIMINARY: PRT: 238 AA.
AC Q91ZG6: 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DM417G6.2.1 (novel protein (isoform 1)).
GN DM417G6.2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wall M.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL358892; CAC42202.1;
SQ SEQUENCE 238 AA: 28038 MW: 8E55135F6EC4E9F9 CRC64;
```

Query Match 13.3%; Score 427; DB 11; Length 238;
Best Local Similarity 47.1%; Pred. No. 1.1e-21;
Matches 82; Conservative 34; Mismatches 54; Indels 4; Gaps 2;

```

QY 398 NKLOKLLILLRGLPGSGKTTLSRILLQGNRDGIVSTDDYFHHODG-YRYNNVNLGDAHD 456
DB 36 HSFRLHYLLRGLPGSGKTTLSRILLQOHYDPRALIFSTDDFFKEDGTTFEFPNLLLEAAHE 95
QY 457 WNQRARAKAIDQGRSPVIIDNTNQAWEMKPYVEVAIGKGYRVEHEPETWKKFDPPEELE 516
DB 96 WNQRARAKAMRGISPIIIDNTNLHAWEMKPYAVNALENENYEVIFREDPTRWKFVQELA 155
QY 517 KRNKUGVSRKKTQAOMLDREYOMSTISVNSVEPSHKST---ORPPPGQGRWR 567
DB 156 RNRHGVPEKTIQRMKERYEHNVTFSVLHAEKPSRANRGNSEPPSSGGYW 209

RESULT 7
QYXZ11 PRELIMINARY; PRT; 969 AA.
AC Q9XZ11
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE BCNA:LD21293 protein.
GN BCNA:LD21293 OR CG7139.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.M.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Butenhoff C.,

```

```

RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacle J.M.,
RA Park S., Sequeira A., Sethi H., Snir E., Svirskas R.R., Weinburg T.,
RA Celniker S.E.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003596; AF51804.1; -
DR EMBL; AF132164; AAD34752.1; -
DR FlyBase; FBgn027532; BCNA:LD21293.
DR InterPro; IPR002625; Smr/Muts2.
DR Pfam; PF01713; Smr; 1.
DR SMART; SM00463; Smr; 1.
KW Hypothetical protein.
SQ SEQUENCE 969 AA; 109357 MW; 5635BCF416BB0AE1 CRC64;

Query Match 10.4%; Score 334; DB 5; Length 969;
Best Local Similarity 31.0%; Pred. No. 1.5e-14;
Matches 84; Conservative 37; Mismatches 92; Indels 58; Gaps 6;

QY 320 NNCVNNNCMTFDQNNETDCSENRSS-----VHPSNGCGCMQDRYVSN 363
DB 33 NNNNNNNNNNNNNNNNNNGCRSGKSGTGRKYATAEAPSSIRMGPSALNSQSNKQFN 92
QY 364 GFCEVRECRKWDCHMDKNGTDRFVNQOFQBEKLNKLLKLLRGLPGSGKTTLSRILL 423
DB 93 SIC---QRAQAGH-----KLMIMRGPSSGSKTSLAESLL 124
QY 424 GQN-----RDGIVSTDDYFHHODGYRYNNVNLGDAHDWNQNRKAOIDQGRSPVII 475
DB 125 RQALLDRHQVDR-FVLSDDYFKTRGYVFNPTLLPAHWNQORVRDKKASGWSPIIV 183
QY 476 DNTNQAWEMKPYVEVAIGKGYRVEHEPETWKKFDPPELEKRNKHGYSRKIAQMDRY 535
DB 184 DNTNTVMWEMQYVQFAVRHGVTLELLEPNTSWCKSASKLACKNVHNPVRENIQRMLERF 243
QY 536 EYQMSISIVMNSVPSH-----KSTQRPPTPP 561
DB 244 ERTTAGELIQLMKETKYSVELPOLRNHPLP 274

RESULT 8
QYXZ11 PRELIMINARY; PRT; 145 AA.
AC Q91ZG7
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE DM41766.2.2 (novel protein (isoform 2)).
GN DM41766.2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wall M.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL358892; CAC42203.1; -
SQ SEQUENCE 145 AA; 17129 MW; 70023679FC802C8C CRC64;

Query Match 8.7%; Score 278.5; DB 11; Length 145;
Best Local Similarity 57.6%; Pred. No. 8.4e-12;
Matches 53; Conservative 15; Mismatches 23; Indels 1; Gaps 1;

QY 398 NKLOKLLILLRGLPGSGKTTLSRILLQGNRDGIVSTDDYFHHODG-YRYNNVNLGDAHD 456
DB 36 HSFRLHYLLRGLPGSGKTTLSRILLQOHYDPRALIFSTDDFFKEDGTTFEFPNLLLEAAHE 95
QY 457 WNQRARAKAIDQGRSPVIIDNTNQAWEMKPY 488
DB 96 WNQRARAKAMRGISPIIIDNTNLHAWEMKPY 127

```

RESULT 9

Q8WTU5 PRELIMINARY; PRT; 173 AA.
ID 08WTU5
AC 08WTU5
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 20.7 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Brain;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC022188; AAH22188.1; -
KW Hypothetical protein.
SQ SEQUENCE 173 AA; 20678 MW; E79510AF67362A1E CRC64;

Query Match
Best Local Similarity 56.5%; Score 272.5; DB 4; Length 173;
Matches 52; Conservative 16; Mismatches 23; Indels 1; Gaps 1;

QY 398 NKLQKLLIRLPGSGKTLRLILGQNRDGIYFTDDYFHODG-YRYNVNQLGDADH 456

DB 38 HSRFRKLYLRLGPGSGKTLRLILGQNRDGIYFTDDYFHODG-YRYNVNQLGDADH 97

QY 457 WNONRAKALDQGRSPYIINTNINQAMWKPY 488

DB 98 WNOFRARAKMRNGISPIIINTNINQAMWKPY 129

RESULT 10

Q9Y273 PRELIMINARY; PRT; 137 AA.

ID 09Y273
AC 09Y273
DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Hypothetical 16.4 kDa protein (92M18.2.1) (Novel protein).
GN 92M18.2.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RA Rhodes S.;

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [1]

RP SEQUENCE FROM N.A.
RA Hunt A.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL049788; CAB42446.1; -

DR EMBL: AL049785; CAB42443.1; -

DR EMBL: AL049786; CAB42444.1; -

DR EMBL: 273359; CAC94789.1; -

KW Hypothetical protein.
SQ SEQUENCE 137 AA; 16393 MW; BC0939347608B77C CRC64;

Query Match
Best Local Similarity 7.7%; Score 247; DB 4; Length 137;
Matches 48; Conservative 21; Mismatches 39; Indels 4; Gaps 1;

QY 469 GRSPVITDNTNINQAMWKPYEVAIGKGRVFEFETWKFDEPELEKRNKHGVSRKI 528

DB 4 GISPIIDNTNINQAMWKPYEVAIGKGRVFEFETWKFDEPELEKRNKHGVSRKI 63

QY 529 AQMLDREYOMGISIVNVSPEPSHKT---HKSTORPPGGRORWGSGLSGSHNR 576

DB 64 HKMKEREHVHTFHSVLAHKRPSRMNRNQRNNALPSSNNARVWNSYTFEPFR 115

RESULT 11

Q8VE50 PRELIMINARY; PRT; 134 AA.

ID 08VE50
AC 08VE50
DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Similar to hypothetical gene CG018.
GN A1428195.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.
RA Strausberg R.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC019754; AAH19754.1; -

DR MGI: 2140872; A1428195.
SQ SEQUENCE 134 AA; 15784 MW; 572C8A1F870CD649 CRC64;

Query Match
Best Local Similarity 44.1%; Score 242.5; DB 11; Length 134;
Matches 45; Conservative 21; Mismatches 33; Indels 3; Gaps 1;

QY 469 GRSPVITDNTNINQAMWKPYEVAIGKGRVFEFETWKFDEPELEKRNKHGVSRKI 528

DB 4 GISPIIDNTNINQAMWKPYEVAIGKGRVFEFETWKFDEPELEKRNKHGVSRKI 63

QY 529 AQMLDREYOMGISIVNVSPEPSHKT---QRPPEGRORW 567

DB 64 QRMKEREHVHTFHSVLAHKRPSRMNRNQRNNALPSSNNARVWNSYTFEPFR 105

RESULT 12
ID 025201 PRELIMINARY; PRT; 232 AA.

ID 025201
AC 025201
DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Hypothetical protein Hp0454.
GN Hp0454.

OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]

RP SEQUENCE FROM N.A.
RA STRAIN=26695 / ATCC 700392;

RC MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervatage A.R., Clayton R.A., Sutton G.G.,

RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,

RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,

RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Melman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,

RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;

RT "The complete genome sequence of the gastric pathogen Helicobacter
pylori."

RL Nature 388:539-547(1997).
DR EMBL: AE000560; AAD07520.1; -

DR TIGR: HP0454; -
KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 232 AA; 27447 MW; 398AB5A07CA27668 CRC64;

Query Match
Best Local Similarity 6.1%; Score 195.5; DB 16; Length 232;
Matches 57; Conservative 33; Mismatches 63; Indels 25; Gaps 8;

```

QY 394 EKLKLNKLOKLLILLRCLPGSKTTLRSL-----LGNQRDGIVFSTDDYFHHQD--GYR 445
DB 8 QKRMKSNKVLINRAIPGGGKSTSLIKQIEELAKSLGHSIS--VHSTDEYFIQTDEEGIR 65
QY 446 YNV--NOLGDAHDWQNRKAQIDQGRSPVITDNTNQAWEMKPYVEVAIGKGYRVEFH 502
DB 66 HYVDKKLNEYHQNQAEKALENRIDIVVCDNTNPNESWQSKPYTDWAREFGYKILLI 125
QY 503 EPTWKKFDPPELEKRNKH--VSR--KKIAQMLDRYEQMSISIVMSNVSPESHKSTQ 556
DB 126 D-----FKNRHLETPMDYGDVQAQICKRPGIAKHVDYFLERVL--VEPDQYEQK 175

RESULT 13
Q19804
ID Q19804 PRELIMINARY; PRT; 162 AA.
AC Q19804;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE Hypothetical 18.4 kDa protein.
GN F26A1.14.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Fulton L.;
RT "The sequence of C. elegans cosmid F26A1.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U27312; AAA68256.1;
KW Hypothetical protein.
SQ SEQUENCE 162 AA; 18438 MW; FC2D9FC15F985140 CRC64;

Query Match 5.1%; Score 164.5; DB 5; Length 162;
Best Local Similarity 26.2%; Pred. No. 0.00063;
Matches 39; Conservative 35; Mismatches 46; Indels 29; Gaps 4;

QY 394 EKLKLNKLOKLL-----ILLRGLPGSKTTLRSLLLGNQRDGIVFSTDDYFHHQDGYR 446
DB 13 DAEISDIQKCLVEGHTILINGVTGSGKSTLARELVNHSNGVIVKND-----VSN 64
QY 447 NVNQLGDAHDWQNRKAQIDQGRSPVITDNTNQAWEMKPYVEVAIGKGYRVEFHEPET 506
DB 65 NITR-----SVRFIDEDKHLIVVDENQVRSSVKKFAELAVNGHVEIFVLEPDT 114
QY 507 WKKFDPPELEKRNKH-----GYSRKKIAQM 531
DB 115 DWRHDAIECKKRSEKDEIGSTESKIMQL 143

RESULT 14
Q9KRM6
ID Q9KRM6 PRELIMINARY; PRT; 149 AA.
AC Q9KRM6;
DT 01-OCT-2000 (TremBLrel. 15, Created)

```

```

DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE Hypothetical protein VC1610.
GN VC1610.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umavam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Olin H., Dragol I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004238; AAF94764.1;
DR TIGR; VC1610;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 149 AA; 17450 MW; 45D56676A4FF8A9A CRC64;

Query Match 5.0%; Score 161; DB 16; Length 149;
Best Local Similarity 29.5%; Pred. No. 0.00098;
Matches 41; Conservative 26; Mismatches 46; Indels 26; Gaps 6;

QY 404 LILLRGLPGSKTTLRSLLLGNQRDGIVFSTDDYFHHQDGYRNVNOLGDAHDWQNR 462
DB 30 LTLIRGLPGSKSTLAKTL-----SAVHLEADYVFNVPQGEYHERPEYLAQAHWCQOQT 84
QY 463 KOAIDQGRSPVITDNTNQAWEMKPYVEVAIGKGYRVEF-----HEPETWKKFDPPELEK 517
DB 85 EYWLQOQKGD--VVVSTNFTVRHWMAMVYRKLA--RQYRAKLTILVCRE-----RY 129
QY 518 RNKGVSRKKIAQMLDRYE 536
DB 130 QNVHGVDEATVERMRQOQW 148

RESULT 15
Q91V42
ID Q91V42 PRELIMINARY; PRT; 400 AA.
AC Q91V42;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Cyclic nucleotide phosphodiesterase 1.
GN CNP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ILS, AND ISS;
RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Canniff J.,
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;
RT "High-Throughput Sequence Identification of Gene Coding Variants
Within Alcohol-Related QTLs.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE ADENYLATE KINASE FAMILY.
DR EMBL; AF332056; AAK56085.1;
DR EMBL; AF332055; AAK56084.1;
DR MGD; MGI:88437; Cnpl.
DR InterPro; IPR001230; Prenyl_site.
DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
KW Kinase; Transferase.
SQ SEQUENCE 400 AA; 44654 MW; B974404499DE29BE CRC64;

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 18:04:15 ; Search time 12.2667 Seconds
(without alignments)
55.005 Million cell updates/sec

Title: US-09-924-654-4_COPY_550_565
Perfect score: 92
Sequence: 1 PSHKSTGRPPPGGRQ 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188354 segs, 42170167 residues

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/1/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubppa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep.*
10: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep.*
12: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep.*
13: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep.*
14: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	583	10 US-09-924-654-4	Sequence 4, Appl1
2	92	100.0	594	10 US-09-925-300-1079	Sequence 1079, Ap
3	49	53.3	901	10 US-09-737-149-33	Sequence 33, Appl
4	49	53.3	998	9 US-10-106-534-2	Sequence 2, Appl1
5	49	53.3	1043	10 US-09-737-149-8	Sequence 8, Appl1
6	48.5	52.7	83	9 US-10-028-072-42	Sequence 42, Appl1
7	48.5	52.7	83	9 US-10-121-049-42	Sequence 42, Appl1
8	48.5	52.7	83	9 US-10-123-904-42	Sequence 42, Appl1
9	48.5	52.7	83	9 US-10-140-470-42	Sequence 42, Appl1
10	48.5	52.7	83	9 US-10-175-746-42	Sequence 42, Appl1
11	48.5	52.7	83	9 US-10-176-918-42	Sequence 42, Appl1
12	48.5	52.7	83	9 US-10-176-921-42	Sequence 42, Appl1
13	48.5	52.7	83	9 US-10-137-865-42	Sequence 42, Appl1
14	48.5	52.7	83	9 US-10-140-474-42	Sequence 42, Appl1
15	48.5	52.7	83	9 US-10-142-431-42	Sequence 42, Appl1
16	48.5	52.7	83	9 US-10-143-114-42	Sequence 42, Appl1
17	48.5	52.7	83	9 US-10-140-002-42	Sequence 42, Appl1
18	47	51.1	196	9 US-09-989-920-224	Sequence 224, App
19	47	51.1	345	9 US-10-036-729-4	Sequence 4, Appl1

20	46	50.0	80	10 US-09-864-761-41757	Sequence 41757, A
21	46	50.0	468	10 US-09-925-297-736	Sequence 736, Appl
22	46	50.0	506	10 US-09-801-368-90	Sequence 90, Appl
23	45	48.9	274	10 US-09-850-887-4	Sequence 4, Appl1
24	45	48.9	322	10 US-09-823-038A-33	Sequence 33, Appl1
25	44.5	48.4	21	10 US-09-873-106B-22	Sequence 22, Appl1
26	44.5	48.4	70	10 US-09-873-106B-11	Sequence 11, Appl1
27	44.5	48.4	351	10 US-09-796-033-6	Sequence 6, Appl1
28	44.5	48.4	351	10 US-09-730-465-6	Sequence 6, Appl1
29	44	47.8	72	10 US-09-864-761-40031	Sequence 40031, A
30	44	47.8	107	9 US-10-178-213-269	Sequence 269, App
31	44	47.8	132	10 US-09-864-761-43644	Sequence 43644, A
32	44	47.8	466	12 US-10-095-492-16	Sequence 16, Appl
33	44	47.8	836	10 US-09-934-323-5	Sequence 5, Appl1
34	44	47.8	1175	10 US-09-771-161A-224	Sequence 224, App
35	44	47.8	1175	10 US-09-771-161A-225	Sequence 225, App
36	44	47.8	1175	10 US-09-771-161A-226	Sequence 226, App
37	44	46.7	117	10 US-09-864-761-34970	Sequence 34970, A
38	43	46.7	390	10 US-09-925-302-625	Sequence 625, App
39	43	46.7	491	9 US-10-217-774-2	Sequence 2, Appl1
40	43	46.7	538	9 US-09-976-740-43	Sequence 43, Appl1
41	43	46.7	538	12 US-10-023-529-43	Sequence 43, Appl1
42	43	46.7	538	12 US-10-023-529-43	Sequence 43, Appl1
43	43	46.7	674	9 US-10-086-464-14	Sequence 14, Appl1
44	43	46.7	1224	9 US-10-217-774-4	Sequence 4, Appl1
45	42	45.7	88	10 US-09-867-550-1482	Sequence 1482, Ap

ALIGNMENTS

RESULT 1
US-09-924-654-4
; Sequence 4, Application US/09924654
; Patent No. US20020146712A1
; GENERAL INFORMATION:
; APPLICANT: Goold, Richard D.
; APPLICANT: Aeridom, Ingrid E.
; APPLICANT: Selhamer, Jeffrey J.
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: TUMOR SUPPRESSOR
; FILE REFERENCE: PC-0049 CIP
; CURRENT APPLICATION NUMBER: US/09/924,654
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020146712A1 496071CD1
US-09-924-654-4

Query Match 100.0%; Score 92; DB 10; Length 583;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSHKSTGRPPPGGRQ 16
DB 550 PSHKSTGRPPPGGRQ 565
|||||

RESULT 2
US-09-925-300-1079
; Sequence 1079, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101


```
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1079
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (430)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1079

Query Match          100.0%; Score 92; DB 10; Length 594;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSHKSTORPPPPQGRQ 16
   :: |||||
Db 561 PSHKSTORPPPPQGRQ 576

RESULT 3
US-09-737-149-33
; Sequence 33, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,544
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 33
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 901
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-737-149-33

Query Match          53.3%; Score 49; DB 10; Length 901;
Best Local Similarity 66.7%; Pred. No. 62;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 STQRP PPQGRQ 16
   :: |||||
Db 401 ASARPPPPQGAQ 412

RESULT 4
US-09-737-149-8
; Sequence 8, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,362
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,544
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 8
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-737-149-8

Query Match          53.3%; Score 49; DB 10; Length 1043;
Best Local Similarity 66.7%; Pred. No. 72;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 STQRP PPQGRQ 16
   :: |||||
Db 401 ASARPPPPQGAQ 412
```

```
US-10-106-534-2
; Sequence 2, Application US/10106534
; Patent No. US20020168668A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 14691, A Human Glutamate Receptor Family
; TITLE OF INVENTION: Member and Uses Therefor
; FILE REFERENCE: MP101-042P1RM
; CURRENT APPLICATION NUMBER: US/10/106,534
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: 60/279,086
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 998
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-106-534-2

Query Match          53.3%; Score 49; DB 9; Length 998;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 STQRP PPQGRQ 16
   :: |||||
Db 395 ASARPPPPQGAQ 406

RESULT 5
US-09-737-149-8
; Sequence 8, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,362
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,544
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 8
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-737-149-8

Query Match          53.3%; Score 49; DB 10; Length 1043;
Best Local Similarity 66.7%; Pred. No. 72;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 STQRP PPQGRQ 16
   :: |||||
Db 401 ASARPPPPQGAQ 412
```

RESULT 6
US-10-028-072-42
Sequence 42, Application US/10028072
Publication No. US20030004311A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Mei-Qiang
APPLICANT: Geriltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang
TITLE OR INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028, 072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561

PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 4 KSTQPPPPGGRQ 16
: : ||||| |||
DB 35 RSPQPPPP-GRQ 46

RESULT 7

US-10-121-049-42
; Sequence 42, Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-049-42

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 4 KSTQPPPPGGRQ 16
: : ||||| |||
DB 35 RSPQPPPP-GRQ 46

RESULT 8

US-10-123-904-42
; Sequence 42, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC54
; CURRENT APPLICATION NUMBER: US/10/123,904
; CURRENT FILING DATE: 2002-04-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-904-42

OY 4 KSTQPPPPGGRQ 16
: : ||||| |||
DB 35 RSPQPPPP-GRQ 46

RESULT 9

US-10-121-049-42
; Sequence 42, Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 4 KSTORPPPPGRO 16
: | | | | | | | | | |
Db 35 RSPQRPPPP-GRO 46

RESULT 9

US-10-140-470-42
; Sequence 42, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C160
CURRENT APPLICATION NUMBER: US/10/140,470
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 42
LENGTH: 83
TYPE: PRT
ORGANISM: Homo Sapien

US-10-140-470-42

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 4 KSTORPPPPGRO 16
: | | | | | | | | | |
Db 35 RSPQRPPPP-GRO 46

RESULT 10

US-10-175-746-42
; Sequence 42, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel

APPLICANT: Matanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C353
CURRENT APPLICATION NUMBER: US/10/175,746
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 42
LENGTH: 83
TYPE: PRT
ORGANISM: Homo Sapien

US-10-175-746-42

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 4 KSTORPPPPGRO 16
: | | | | | | | | | |
Db 35 RSPQRPPPP-GRO 46

RESULT 11

US-10-176-918-42
; Sequence 42, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C382
CURRENT APPLICATION NUMBER: US/10/176,918
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 42
LENGTH: 83
TYPE: PRT
ORGANISM: Homo Sapien

US-10-176-918-42

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 4 KSTORPPPPGRO 16
: | | | | | | | | | |
Db 35 RSPQRPPPP-GRO 46

RESULT 12

US-10-176-921-42
; Sequence 42, Application US/10176921
; Publication No. US20030027276A1

```
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C286
; CURRENT APPLICATION NUMBER: US/10/176,921
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-921-42

Query Match      52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 4 KSTQRP PPPGGRQ 16
Db 35 RSPQRP PPP-GRQ 46

RESULT 13
US-10-137-865-42
; Sequence 42, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C154
; CURRENT APPLICATION NUMBER: US/10/137,865
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
```

```
US-10-137-865-42

Query Match      52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 4 KSTQRP PPPGGRQ 16
Db 35 RSPQRP PPP-GRQ 46

RESULT 14
US-10-140-474-42
; Sequence 42, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-474-42

Query Match      52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 4 KSTQRP PPPGGRQ 16
Db 35 RSPQRP PPP-GRQ 46

RESULT 15
US-10-142-431-42
; Sequence 42, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
```

APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS
FILE REFERENCE: P3330R1C251
CURRENT APPLICATION NUMBER: US/10/142,431
CURRENT FILING DATE: 2002-05-10
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 42
LENGTH: 83
TYPE: PRT
ORGANISM: Homo Sapien
US-10-142-431-42

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 4 KSTORPPPPGGRQ 16
: ||||| |||
DB 35 RSPQRPPPP-GRQ 46

Search completed: March 10, 2003, 18:19:49
Job time : 12.2667 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 17:46:43 ; Search time 18 Seconds
(without alignments)
1343.372 Million cell updates/sec

Title: US-09-924-654-4

Perfect score: 3209
Sequence: 1 MSYGEIEGKFLGPREETSE.....RORWGSIGSHNRVCVTNNH 583

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	154	4.8	400	1 CN37_BOVIN	P06623 bos taurus
2	148.5	4.6	421	1 CN37_HUMAN	P09543 homo sapien
3	147	4.6	420	1 CN37_MOUSE	P16330 mus musculu
4	131	4.1	1386	1 ZAP3_MOUSE	O99017 mus musculu
5	128.5	4.0	1822	1 ZAP3_HUMAN	P49750 homo sapien
6	128	4.0	964	1 YOY1_CAEL	O09560 caenorhabdi
7	127	4.0	275	1 FRA1_RAT	P10158 rattus norv
8	124	3.9	407	1 YZRS_CAEL	O19683 caenorhabdi
9	121.5	3.8	1059	1 CERD_RAT	P13635 rattus norv
10	119	3.7	1790	1 USOI_YEAST	P25386 saccharomyc
11	118.5	3.7	301	1 KIPN_BPT4	P06855 bacterioph
12	118	3.7	273	1 FRA1_MOUSE	P48755 mus musculu
13	117.5	3.7	952	1 IF41_YEAST	P39935 saccharomyc
14	117	3.6	1658	1 YME7_YEAST	O03661 saccharomyc
15	116	3.6	404	1 CN37_RAT	P13233 rattus norv
16	115	3.6	1164	1 KELL_YEAST	P38853 saccharomyc
17	113	3.5	646	1 CG11_CANAL	P24866 candida alb
18	113	3.5	1157	1 SRA4_HUMAN	O95104 homo sapien
19	111	3.5	1484	1 CES2_HUMAN	O9bxt3 homo sapien
20	110.5	3.4	943	1 ARS2_DROME	O9v9x7 drosophila
21	110.5	3.4	1076	1 YEM3_YEAST	P40021 saccharomyc
22	110.5	3.4	1176	1 YOHB_YEAST	O08236 saccharomyc
23	110.5	3.4	1355	1 SALM_DROME	P39770 drosophila
24	110.5	3.4	1818	1 HMM2_MYCPN	P75471 mycoplasma
25	110	3.4	618	1 MM24_MOUSE	O97052 mus musculu
26	110	3.4	645	1 MM24_HUMAN	O9y5t2 homo sapien
27	110	3.4	1032	1 KINN_HUMAN	O12840 homo sapien
28	109.5	3.4	1147	1 CGAL_HELPY	P00200 helicobacte
29	109	3.4	2774	1 MAPA_RAT	P49266 rattus norv
30	108.5	3.4	271	1 FRA1_HUMAN	P15407 homo sapien
31	108	3.4	618	1 MM24_RAT	O99p66 rattus norv
32	107.5	3.3	710	1 L778_ARATH	O06738 arabidopsis
33	107.5	3.3	712	1 Y352_HUMAN	O15060 homo sapien

34	107.5	3.3	1399	1 Z291_HUMAN	O9b12 homo sapien
35	107.5	3.3	1935	1 MYSS_CYPCA	O90339 cyprinus ca
36	107.5	3.3	1939	1 MYH6_HUMAN	P13533 homo sapien
37	107.5	3.3	2869	1 RBP1_PLAVB	O00798 plasmodium
38	107	3.3	462	1 CATC_MOUSE	P97821 mus musculu
39	107	3.3	725	1 ADDB_RAT	O05764 rattus norv
40	107	3.3	731	1 BAF1_YEAST	P14164 saccharomyc
41	107	3.3	1182	1 CGA2_HELPY	P55746 helicobacte
42	107	3.3	1447	1 BUD4_YEAST	P47136 saccharomyc
43	106.5	3.3	1186	1 CAGA_HELPY	P55980 helicobacte
44	106.5	3.3	1411	1 YM42_YEAST	O03214 saccharomyc
45	106	3.3	506	1 NPL3_HUMAN	O99457 homo sapien

ALIGNMENTS

```
RESULT 1
ID CN37_BOVIN STANDARD; PRT; 400 AA.
AC P06623;
DT 01-JAN-1988 (Rel. 06, Created)
DD 01-JUL-1993 (Rel. 26, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 2',3'-cyclic nucleotide 3'-phosphodiesterase (EC 3.1.4.37) (CNP)
DE (CNPase).
GN CNP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE=87137604; PubMed=3029107;
RA Kurinara T., Fowler A.V., Takahashi Y.;
RT "cDNA cloning and amino acid sequence of bovine brain 2',3'-cyclic-
RL nucleotide 3'-phosphodiesterase."
RT J. Biol. Chem. 262:3256-3261(1987).
RN [2]
RP REVISION TO 318.
RC TISSUE-Brain;
RA Kurinara T., Fowler A.V., Takahashi Y.;
RT "Nucleotide sequence of bovine retina 2',3'-cyclic nucleotide 3'-
RL phosphodiesterase."
RL Nucleic Acids Res. 15:7204-7204(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=87247281; PubMed=3036592;
RA Vogel U.S., Thompson R.J.;
RT "Molecular cloning of the myelin specific enzyme 2',3'-cyclic-
RL nucleotide 3'-phosphodiesterase."
RL FEBS Lett. 218:261-265(1987).
RN [5]
RP CATALYTIC ACTIVITY: Nucleoside 2',3'-cyclic phosphate + H(2)O =
RL nucleoside 2'-phosphate.
RN [6]
RP SUBCELLULAR LOCATION: FIRMLY BOUND TO MEMBRANE STRUCTURES OF BRAIN
WHITE MATTER.
RN [7]
RP MET-1 MAY BE REMOVED AFTER TRANSLATION.
RN [8]
RP THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; J02659; AAA30456.1; ALT_SEQ.
DR EMBL; Y00405; CA68466.1;
DR EMBL; M27606; AAA30457.1;
DR PIR; A26861; ESBOP3.
DR InterPro; IPR001230; Prenyl_site.
KW Hydrolase; Membrane; Brain.
SQ SEQUENCE 400 AA: 44875 MW: 3448FC367D647CF8 CRC64;

Query Match 4.8%; Score 154; DB 1; Length 400;
Best Local Similarity 28.6%; Pred. No. 0.0092;
Matches 44; Conservative 37; Mismatches 49; Indels 24; Gaps 5;

QY 394 EKLKQLQ--KLILLRGLPGSGKTTLSRILLGQNRDGIIVFTDDYFHHQDGYRYNVNQL 451
DB 20 ETVATLQCKFLFILRGLPGSGKSTLAFIVDKTRDGYKMSADYKKTTPGARGSFSE- 78
QY 452 GDAHDNQNRKAOIDQGRSP-----VIIDNTNIQAWEMKPYVEVAIGKGYRVEFHE 503
DB 79 ----EYKQ-----LDELAACCRDRFVLVLDNTNHERERLEQLFELADQYQYQVVLVE 128
QY 504 PETWKFDEPEELEKRNKHGVSR---KKIAQMLDR 534
DB 129 PKTAWRLDCAQLKERNQWLSADLUKLPGLK 162

RESULT 2
CN37_HUMAN
ID CN37_MOUSE STANDARD; PRT; 421 AA.
AC P09543;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 2',3'-cyclic nucleotide 3'-phosphodiesterase (EC 3.1.4.37) (CNP)
DE (CNPase).
GN CNP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93050745; PubMed=1385234;
RA Thompson R.J.;
RT "2',3'-cyclic nucleotide-3'-phosphohydrolase and signal transduction
RT in central nervous system myelin.";
RL Biochem. Soc. Trans. 20:621-626(1992).
RN [2]
RP SEQUENCE FROM N.A. (SHORT FORM).
RC TISSUE=Brain;
RX MEDLINE=88209067; PubMed=2835044;
RA Kurihara T., Takahashi Y., Nishiyama A., Kumanishi T.;
RT "cDNA cloning and amino acid sequence of human brain 2',3'-cyclic-
RT nucleotide 3'-phosphodiesterase.";
RL Biochem. Biophys. Res. Commun. 152:837-842(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93314977; PubMed=8392017;
RA Monoh K., Kurihara T., Takahashi Y., Ichikawa T., Kumanishi T.,
RA Hayashi S., Minoshima S., Shimizu N.;
RT "Structure, expression and chromosomal localization of the gene
RT encoding human 2',3'-cyclic nucleotide 3'-phosphodiesterase.";
RL Gene 129:297-301(1993).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93080285; PubMed=1360194;
RA Douglas A.J., Fox M.F., Abbott C.M., Hinks L.J., Sharpe G.,
RA Povey S., Thompson R.J.;
RT "Structure and chromosomal localization of the human 2',3'-cyclic
RT nucleotide 3'-phosphodiesterase gene.";
RL Ann. Hum. Genet. 56:243-254(1992).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
```

```
RC TISSUE=Brain, and Skin;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Nucleoside 2',3'-cyclic phosphate + H(2)O -
CC nucleoside 2'-phosphate.
CC -!- SUBCELLULAR LOCATION: FIRMLY BOUND TO MEMBRANE STRUCTURES OF BRAIN
CC WHITE MATTER.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: CNP1/DNAI AND CNP1/DNAI
CC (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; S46849; AAB23928.2;
DR EMBL; S46843; AAB23928.2; JOINED.
DR EMBL; S46845; AAB23928.2; JOINED.
DR EMBL; S46846; AAB23928.2; JOINED.
DR EMBL; M19650; AAA35704.1;
DR EMBL; D13146; BRA39694.1;
DR EMBL; D13144; BRA39694.1; JOINED.
DR EMBL; D13145; BRA39694.1; JOINED.
DR EMBL; D13146; BRA02435.1;
DR EMBL; D13144; BRA02435.1; JOINED.
DR EMBL; D13145; BRA02435.1; JOINED.
DR EMBL; S50017; AAB24298.2;
DR EMBL; S50013; AAB24298.2; JOINED.
DR EMBL; S50014; AAB24298.2; JOINED.
DR EMBL; S50016; AAB24298.2; JOINED.
DR EMBL; BC001362; AAH01362.1;
DR EMBL; BC006392; AAH06392.1;
DR EMBL; BC011046; AAH11046.1;
DR EMBL; BC028040; AAH28040.1;
DR PIR; A27703; A27703.
DR PIR; JCI518; JCI518.
DR Genew; HGNC:2158; CNP.
DR MIN; 123830;
DR InterPro; IPR001230; Prenyl_site.
KW Hydrolase; Membrane; Brain; Alternative splicing.
FT VARSPPLIC 1 20
SQ SEQUENCE 421 AA: 47578 MW: CA6D0097DFD87255 CRC64;

Query Match 4.6%; Score 148.5; DB 1; Length 421;
Best Local Similarity 28.7%; Pred. No. 0.022;
Matches 39; Conservative 36; Mismatches 56; Indels 5; Gaps 2;

QY 402 KLLILLRGLPGSGKTTLSRILLGQNRDGIIVFTDDYFHHQDGYRYNVNQLGDAHDNQNR 461
DB 50 KTLFILRGLPGSGKSTLAFIVDKTRDGYKMSADYKKTTPGARGAFSE--EYKRLDEL 107
QY 462 AKQAIQGRSPVIIDNTNIQAWEMKPYVEVAIGKGYRVEFHEPETWKFDEPEELEKRNKH 521
DB 108 AAYCRRDRIRILVLDNTNHERERLEQLFELADQYQYQVVLVEPKTAWRLDCAQLKKNOW 167
QY 522 GVSR---KKIAQMLDR 534;
DB 168 QLSADDLKKLPGLK 183

RESULT 3
CN37_MOUSE
ID CN37_MOUSE STANDARD; PRT; 420 AA.
AC P16330;
DT 01-AUG-1990 (Rel. 15, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 2',3'-cyclic nucleotide 3'-phosphodiesterase (EC 3.1.4.37) (CNP)
DE (CNPase).
GN CNP OR CNP1.
```


OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM CNPI).
RX MEDLINE=90121227; Pubmed=2558653;
RA Monoh K., Kurihara T., Sakimura K., Takahashi Y.;
RT "Structure of mouse 2',3'-cyclic-nucleotide 3'-phosphodiesterase
RT gene";
RL Biochem. Biophys. Res. Commun. 165:1213-1220(1989).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS CNPI AND CNPII).
RX STRAIN=DA; MEDLINE=90358801; Pubmed=2167669;
RA Kurihara T., Monoh K., Sakimura K., Takahashi Y.;
RT "Alternative splicing of mouse brain 2',3'-cyclic-nucleotide 3'-
RT phosphodiesterase mRNA";
RL Biochem. Biophys. Res. Commun. 170:1074-1081(1990).
CC -1- CATALYTIC ACTIVITY: Nucleoside 2',3'-cyclic phosphate + H2O =
CC nucleoside 2'-phosphate.
CC -1- SUBCELLULAR LOCATION: FIRMLY BOUND TO MEMBRANE STRUCTURES OF BRAIN
CC WHITE MATTER.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: CNPI/DNAI AND CNPII/DNAII
CC (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M31810; AAA37429.1; -
DR EMBL: M58045; AAA37430.1; -
DR EMBL: D38642; BAA07621.1; JOINED.
DR EMBL: D38640; BAA07621.1; JOINED.
DR EMBL: D38641; BAA07621.1; JOINED.
DR EMBL: D38642; BAA07622.1; -
DR EMBL: D38640; BAA07622.1; JOINED.
DR EMBL: D38641; BAA07622.1; JOINED.
DR PIR: A35708; ESM332.
DR MGI: MGI:88437; Cnpi.
DR InterPro: IPR001230; Prenyl-site.
KW Hydrolyase; Membrane; Brain; Alternative splicing.
FT VAAPSPIC 1 20 MISSING (IN ISOFORM CNPI).
FT CONFLICT 115 115 M -> I (IN REF. 1).
FT CONFLICT 136 136 L -> Y (IN REF. 1).
SQ SEQUENCE 420 AA; 47123 MW; DD96FEED47AD15D3 CRC64;
Query Match 4.6%; Score 147; DB 1; Length 420;
Best Local Similarity 27.8%; Pred. No. 0.027;
Matches 40; Conservative 34; Mismatches 48; Indels 22; Gaps 3;
QY 402 KLLILRLPSSGKTTLSRIILGONRGIVSTDDYFHNDGGRYNNQLDADHANNOR 461
DB 50 KTLRLRLPSSGKSTLRLILEKYHDTKMSADAYKIIPSGRDFSE----- 98
QY 462 AKQALDQGRS-----PVIIDNTNIQAMKPRVEVALIKGYRVEHEDETKMKPPE 513
DB 99 AYKRLDELALAYCRRDMFVLVDLTNHERRLDOLFEALDOYOVVLEFKTAVRLDCA 158
QY 514 ELEKRNKHGVR--KRIQMIDR 534
DB 159 QLKERNQWLSADLKLKPGLEK 182
RESULT 4
ZAP3_MOUSE STANDARD; PRT; 1386 AA.
AC O9R017;
DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nuclear protein ZAP3.
GN ZAP3 OR ZAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal liver;
RA Misawa K., Nosaka T., Kitamura T.;
RT "A huge nuclear protein rich in proline similar to human hypothetical
RT protein zap3 and zap113";
RL submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AB033168; BAA85182.1; -
DR MGI: MGI:1926195; Zap3.
KW Nuclear protein.
FT DOMAIN 15 204 PRO-RICH.
FT DOMAIN 355 473 GLN-RICH.
FT DOMAIN 925 1012 ARG-RICH.
SQ SEQUENCE 1386 AA; 155130 MW; D862FE918ED221DF CRC64;
Query Match 4.1%; Score 131; DB 1; Length 1386;
Best Local Similarity 17.1%; Pred. No. 1.2;
Matches 121; Conservative 78; Mismatches 179; Indels 328; Gaps 28;
QY 91 PDVIESIDSOVLQARPPVLSADEIYTSKAFIPIYKPEKKRN-----EG 139
DB 612 PPAAGSONSQIPKPRQALLPTPVSGSTPPS---PYHPPOSQVNSKPLNKVFSSBOG 668
QY 140 RNEAHVLN-----GI-----NRCGGQKQKQKFNSEKSEIDNELFQ 174
DB 669 LGESSALSQSTIAAKDTPVKSGLLADPPKGSFLDGPRGPREQR----- 712
QY 175 FYKEIELEKEKDEGENSCKESEPSOEQFYPFEGHNNGLKLPDEKKDLSNKM-PSHC 233
DB 713 -----EQLOKTKD-----FGSEPOMDHLP-----PPDSRLQNPSPRGVPPPG 751
QY 234 DYQ--QNLGNPDYKPCNGQVITPFCDTSTFSF---RPMQSVYPIYVYPPPLPSLNYH 288
DB 752 SYRPPPMGKPP-----GSIVRPSAPPARSSIPMTRP-----PVPIPPPPPPP 795
QY 289 LNIQFSGPPRPNINPOADDQIQNGYVNNCHVMNMTFPONNETDGCSENR--- 344
DB 796 -----PPPPPPVYKSKTSYVKQERKWDSEFGLM-----DTNDQGLNSEKKRDTA 842
QY 345 -----SSVHPSG----- 351
DB 843 TTPSAPVLPPEPVVSSITPPPGPMGMPMKPPVQHTVDYGHGRDPTKVKVQIPIYE 902
QY 352 -----NGCSMDQRY-----VSN-----GCEVRE----- 370
DB 903 RITLRPPLPPESTFDADAGORDRYDRDRREYFDPKPSNITDHRDKRRETHRRDR 962
QY 371 -----RCWKDHCMKHNGTDR-----FV 388
DB 963 DRYLDYERDRDRERRPRPDNOSYRDKKDHSSSRGSGFDRPSTDKRSNDRPYGPPMFG 1022
QY 389 NQO--FOEKL-----NKLOKLLILRLGL 410
DB 1023 GERTYPERNRPLAPALGHQPPPVPRYEKKPEKKNVDILKPPGRSRRPRIYVIMKGL 1082

QY 411 PGSGKTTLSRILLGNRD-----GIVFSTDDYF-----HHQD-----GYRN 447
||||| : : : : : ||||| : : : : :
Db 1083 PGSGKTHVAKLRDKEVEFGGPAPRVLSDDYFIAEVEKEEDPDGSKKKVMEYEYE 1142
QY 448 VNLGDAHDWNRQAQADQGRSP-VIIDNTNIOAWMKPYVEVAIGKGYRVEFHEPET 506
| : : : : : | : : : : : | : : : : :
Db 1143 ADMEETRYTSMETFKTKLLDDGFFFIILDAINDRVHFDQFWSAKTKGFVYLAESA 1202
QY 507 WKKFDPPELEKKNKGVSRRKIAQMLDRYEQMSISIVMNSVEPSH 552
| : : : : : | : : : : : | : : : : :
Db 1203 ----DNQTCGRNIGRKLKINKNAEHW-----VAPRH 1233

RESULT 5

ZAP3_HUMAN
ID ZAP3_HUMAN STANDARD; PRT: 1822 AA.
AC P49750; P49752; O9P1V7;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nuclear protein ZAP3 (ZAP113).
OS ZAP3.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rowen L., Madan A., Qin S., Abbasi N., Baradarani L., Birditt B.,
RA Bloom S., Dors M., Dickhoff R., Fleetwood P., Harrison G., James R.,
RA Kaur A., Madan A., Owen M.P., Ratcliffe A., Shafer T., Hood L.,
RA "Sequencing of human chromosome 14q24.3 region."
RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE OF 539-847 AND 1397-1822 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95319502; PubMed=7596406;
RA Sherrington R., Rogaeve E.I., Liang Y., Rogaeve E.A., Levesque G.,
RA Ikeda M., Chi H., Lin C., Li G., Holman K., Tsuda T., Mar L.,
RA Foncin J.-F., Bruni A.C., Montesi M.P., Sorbi S., Rainero I.,
RA Pinassi L., Nee L., Chumakov I., Pollen D., Brookes A.,
RA Sanseau P., Polinsky R.J., Wasco W., da Silva H.A.R., Haines J.L.,
RA Pericak-Vance M.A., Tanzi R.E., Roses A.D., Fraser P.E.,
RA Rommens J.M., St George-Hyslop P.H.;
RT "Cloning of a gene bearing missense mutations in early-onset familial
RT Alzheimer's disease."
RL Nature 375:754-760(1995).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 1661.
CC [2]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC [1]

DR EMBL; AC007956; AAF61275.1;
DR EMBL; L40403; AAC42008.1; ALT_FRAME.
DR EMBL; L40400; AAC42006.1;
KW Nuclear protein.
FT DOMAIN 15 205 PRO-RICH.
FT DOMAIN 382 430 GIN-RICH.
FT DOMAIN 807 1209 ARG-RICH.
FT DOMAIN 1488 1577 ARG-RICH.
FT CONFLICT 621 621 P -> S (IN REF. 2).
FT CONFLICT 1404 1404 T -> I (IN REF. 2).
FT CONFLICT 1821 1821 K -> E (IN REF. 2).
SQ SEQUENCE 1822 AA; 204947 MW; 8E6CB83FE540C7D2 CRC64;

Query Match

4.0%; Score 128.5; DB 1; Length 1822;

Best Local Similarity 17.1%; Pred.No. 2.5;
Matches 96; Conservative 63; Mismatches 134; Indels 267; Gaps 20;
QY 171 ELFOFYKEIEEKEKOGFENSCKESEPQEOVFYFECHNNGLLKPDEKKDLNKNKAMP 230
|| : : : : : || : : : : :
Db 1267 ELKMLREQEQLQKKD-----FGSEQMAHLP-----PQESR----- 1300
QY 231 SHCDYQOQNGNPDKYPONGQVIP-----TFCDTSFTSFRPEWOSVVPF 274
| : : : : : | : : : : : | : : : : :
Db 1301 ----LQNTSSRPGMYPPPGSYRPPPMGKPPGSIVRSPAPRSPVTRP-----PV 1349
QY 275 IVPYGPPLPSNLHNIQRFSGPPNPPSIFQADDSIQNGYVYNNCHVNNCHMTFDO- 333
| : : : : : | : : : : : | : : : : :
Db 1350 PIPPPPPPPPL-----PPPPP--VIKPTSAVEQERWEDSFYGLWD--TNDEQ 1394
QY 334 --NNEY-----TDCS 341
| : : : : : | : : : : :
Db 1395 GLNSEKSETAIPASAPVLPVPPVHSSIPPVPGVPMGMPMSPKPPVQOTVDYGHGRDIS 1454
QY 342 ENRSVHFSGNGCSM-----QDRY-----VSNFGCEVRE- 370
| : : : : : | : : : : : | : : : : :
Db 1455 TNKVEQIPYGERITLRPDPLPERSTFETEHAGORDRYDRERDEPFYDRQSNVIADHRDF 1514
QY 371 -----RCWKDHCMKXKNGTDR----- 386
Db 1515 KRDRTHRDGRGVIDYDRDRFRDRPRDRRAQSYRDKKDHSSRRGGFDRPSYDRK 1574
QY 387 -----FVNOQ--FQBEKL----- 397
Db 1575 SDRPVYEGSMFGGRRRYPEERMLPAPLSHQPPAPRVEKKPKESKNVDILKPPGRE 1634
QY 398 NKLOKLLILLRGLPGSGKTTLSRILLQGNRQDGVFTDDYFHHQDGYRYNNVQLGDAHDW 457
| : : : : : | : : : : : | : : : : :
Db 1635 SRPERIVVIMRGLPGSGKTHVAKLR-----VMEYEYEAEMEETRYTSMFK----- 1680
QY 458 NONRAKQAIQGRSP-VIIDNTNIOAWMKPYVEVAIGKGYRVEFHEPETWKKFDPPELE 516
Db 1681 ----TFKTLDDGFFFIILDAINDRVHFDQFWSAKTKGFVYLAESA-----DNQTCG 1733
QY 517 KRNGKGVSRKKIAQMLDRYE 536
||| || : : : : : | : : : : :
Db 1734 KRNIHGRKLKINKNAEHW 1753
RESULT 6
YQYL_CAEEL STANDARD; PRT: 964 AA.
ID YQYL_CAEEL
AC Q09560;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 108.9 kDa protein F36G3.1 in chromosome X.
GN F36G3.1
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA White S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Jones S.J.M.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC [1]

or send an email to license@slb.stb.ch.

CC EMBL; Z47069; CAAB738.1; -
 DR WormPep; F36G3.1; CE15979.
 KM Hypothetical protein; Coiled coil.
 FT DOMAIN 656 840 COILED COIL (POTENTIAL).
 SQ SEQUENCE 964 AA; 108868 MW; DABA6BE59350F076 CRC64;

Query Match 4.08; Score 128; DB 1; Length 964;
 Best Local Similarity 20.5%; Pred. No. 1.2;
 Matches 96; Conservative 76; Mismatches 171; Indels 126; Gaps 23;

QY 2 SYGIE-CKRLGPREETSPRCKKLKSTESYVFHNHNSNADFRIDEKGNMVPYTI 60
 DB 227 SFGSSDGDHLEVKSEATNPE---YNTSEF-----EVKNDIVYENV 268
 QY 61 DVRGHSYLENKIKTT--DLHRPLHDEMPGNRPVIESIDSVLOEARPLVSADEIYS 118
 DB 269 N-----ESTIDSVVIDSGTPTKDKR--NEVNISQNVNVLRLSH--LENYDDE--T 314
 QY 119 TSKAFIPKPKPKPKRNEGRNEAHVNLGINDRGQKE--KQFNSKSEIDNLFQFY 176
 DB 315 TMTVIACVYDAQENKCE--KNEEDLSPYLDENKIKNSPEKFCFFRAE----SECY 367
 QY 177 KEIELEKREKDFENSKESPEQOFVPEYEGHNNGLKDEEKNLSNAMPSHDQY 236
 DB 368 EQETETDQIK-----QFVPLME-----VSANOEFDELCDIE 399
 QY 237 QNLNEDPKYPCNGQVIFTECDTSFTSFRPEW--QSVYPTVPYGPPLPSNYHLNIQRF 294
 DB 400 RNGNSNDKDR-----PKDLNENKAPDAEHENSYLEY-----EQLPVFSSETIAKT 446
 QY 295 SGPNPSPNIFQADDSQ-----IQNGY--YVNCCHVMNCTEDQNNETDCSENK 344
 DB 447 PIPVHEHSISGVSKEQSVFNQAVQNELMPTHTIISDEVTSTADNNIPFSDCSTHK 506
 QY 345 SSV-----HPSNGSGMODRYNSNCFCEYRECKMDHCKDKH-----NCGD- 385
 DB 507 VKVDFDGSIGKEEELSSSTSSSEF-EFAEPVTENHIODERAQALRNSIIRYPSNDTD 565
 QY 386 ---BFVNOQFOEKL-----NKLQKLILRLGPGSGKTLSPILLQ 425
 DB 566 DELDSVGDFEDDLAVQVIAEVEQLVAALNAGRDEQMSAVMWK 614

RESULT 7
 FRAL_RAT STANDARD; PRT; 275 AA.
 ID P10158;
 AC 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fos-related antigen 1 (FRA-1).
 GN FOSL1 OR FRA1.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88261282; Pubmed=1133553;
 RA Cohen D.R., Curran T.;
 RT "fra-1: a serum-inducible, cellular immediate-early gene that encodes
 a fos-related antigen";
 RL Mol. Cell. Biol. 8:2063-2069(1988).
 RN 121
 RP SEQUENCE OF 1-43 FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA MEDLINE=95311973; Pubmed=7791782;
 RA Bergers G., Granger P., Braselmann S., Wrighton C., Busslinger M.;
 RT "Transcriptional activation of the fra-1 gene by AP-1 is mediated by
 RT regulatory sequences in the first intron";
 RL Mol. Cell. Biol. 15:3748-3756(1995).

CC -1- SUBUNIT: HETERODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- INDUCTION: SERUM-INDUCIBLE.
 CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. FOS SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-stb.ch/announce/>
 CC or send an email to license@slb-stb.ch).

CC EMBL; M19651; AAA4171.1; -
 DR EMBL; U24154; AAA8204.1; -
 DR PIR; A27722; TVRTER.
 DR HSSP; P01100; 1FOS.
 DR TRANSFAC; T01208;
 DR InterPro; IPR000837; Leuzip_Fos.
 DR InterPro; IPR004827; TF_bzip.
 DR Pfam; PF00170; bzip.1.
 DR PRINTS; PR00042; LEUZIPRPOS.
 DR SMART; SM00338; BRLZ.1.
 DR PROSITE; PS00036; BZIP_BASIC.1.
 KM Nuclear protein; DNA-binding.
 FT DOMAIN 105 108 POLY-GLU.
 FT DVA_BIND 113 131 BASIC MOTIF.
 FT DOMAIN 135 163 LEUCINE-ZIPPER.
 FT DOMAIN 255 259 POLY-SER.
 SQ SEQUENCE 275 AA; 103726AD5D1FAB2F CRC64;

Query Match 4.08; Score 127; DB 1; Length 275;
 Best Local Similarity 22.9%; Pred. No. 0.28;
 Matches 62; Conservative 30; Mismatches 91; Indels 88; Gaps 13;

QY 90 RPDVIESIDSVLOEAR---PPL--VSADDEI-YTSKAFIGPI-----YRP- 130
 DB 20 RPAQPOQAQOTVYQOQFHLVPSINAVSGSELQMMVQPHLGSGVPRPLTYQYSPQ 79
 QY 131 -----PEKKRNEGRNEAHVNLGINDRGQKEKQKF-NSEK 165
 DB 80 PRPGVIRALGPPRRRRRRCQISPEEERRRVRERRKLAACRNRRKELTDFLQAE 139
 QY 166 SEINNELFQFKTELEKEKDFENSKESPEQOQFVPEYEGHNNGLKPDSEKDL 225
 DB 140 DKLEDEKSGLORETELEQKQERLE-----LVLEAHRPLCKIPEEDKDTG 185
 QY 226 NKAMPSHCDYQONLGNEDPKYPCN-----GQVI-----PTFCDT-SFTSFRPEW 268
 DB 186 GTSSTS-----GAGSPRG--PCRVPICISLSPGVLEPALHTLTPTSLTPFPSTL 237
 QY 269 QSVYPTVPYGPPLPSNYHLNIQRFSGPPN 299
 DB 238 VFTYF-----STPEPCSSAHRKSSSSGSDPS 263

RESULT 8
 YZRS_CAEEL STANDARD; PRT; 407 AA.
 ID YZRS_CAEEL
 AC 019683; 09G064;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein F21D5.5 in chromosome IV.
 GN F21D5.5.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Berks M.;

```

RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Kohara Y., Shin'ichi T., Suzuki Y., Sugano S., Potdevin M.,
  Thierry-Mieg Y., Thierry-Mieg D., Thierry-Mieg J.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO YEAST YMR156C, S.POMBE SPAC23C11.04C AND
  ACNVP ORE33.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z54271; CAA91035.2; -
DR EMBL; AF292046; AAG41142.1; -
DR WormPep; F21D5.5; CE27135.
KW Hypothetical protein.
SQ SEQUENCE 407 AA; 46188 MW; A17DCAFCB979EF3C CRC64;
Query Match 3.9%; Score 124; DB 1; Length 407;
Best Local Similarity 19.8%; Pred. No. 0.71;
Matches 78; Conservative 55; Mismatches 130; Indels 130; Gaps 16;
QY 182 LSEKDGFGNSCKESEPSEQEVPF-----YEGHNNGLLKPDEKKDLNKNKAMPCHCDVQ 236
DB 1 MKRAMDGAKEAKKSDTDLFGRPLKKTGSWESKNGDLMIFTH-----SECEGK 50
QY 237 ONLGNPEKPYPCNGOVIPTFCDTSTSPFQWQSYPTVPYGPPLPSLNYHL----- 289
DB 51 EKIA----AFDMGLIKTKSKVFTPCNQDWLYDSTSPDFKLLHSDGFKIVFTNQK 106
QY 290 -----NIQFSGPPNPSPNIFQADDSQIQNGYVNNCHVNNCMFTFDQNN 335
DB 107 GIHAGKVDNREPKKIEATVGLGIPVQAFV-----VAGHYRKPVCVGMNELKL--RN 159
QY 336 EYTDSENRSVHPGNSGCMODRIVSN--GFCVERCKWCHCKMDKNGTDRF-----VN 389
DB 160 DEVEINEKESI-----FVGDAAGRIKTTSPKKD-----HSYADRFFAANGV 201
QY 390 QOQF-----EELKNLKLKLLILLRLPGS 413
DB 202 VKFQTEPEFFGSKVDPEWGPFPNFKPLFSEITELEPHDAQLSKSEKILLMVGFPQS 261
QY 414 GKTTLIRLLGNRDGIVFSTDYDTHHQDGYR-YVNNQLGDADHNNQNR--KQAIQDGR 470
DB 262 GKSTFAKML-----GHQHDYKIVNRDTIG---TWQKCVAAFRSLYADCK 302
QY 471 SPVIDNTNIQAWEMKPYVEVAIGKGYRVEFHE 503
DB 303 S-VWIDNTSPDLRSKRYIDVAKELGVPICRFE 334
RESULT 9
CERU_RAT
ID CERU_RAT STANDARD: PRT; 1059 AA.
AC P13635; Q64719;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ceruloplasmin precursor (EC 1.16.3.1) (ferroxidase).
GN CP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE=Lung, and Liver;
RA MEDLINE=90237081; PubMed=2332446;
RA Fleming R.E., Gitlin J.D.;
RT "Primary structure of rat ceruloplasmin and analysis of
  tissue-specific gene expression during development.";
RL J. Biol. Chem. 265:7701-7707(1990).
RN [2]
RP SEQUENCE OF 257-294; 571-612 AND 823-892 FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=87137545; PubMed=3818625;
RA Aldred A.R., Grimes A., Schreiber G., Mercer J.F.B.;
RT "Rat ceruloplasmin. Molecular cloning and gene expression in liver,
  choroid plexus, yolk sac, placenta, and testis.";
RL J. Biol. Chem. 262:2875-2878(1987).
CC -!- FUNCTION: CERULOPLASMIN IS A BLUE, COPPER-BINDING (6-7 ATOMS PER
  MOLECULE) GLYCOPROTEIN FOUND IN PLASMA. FOUR POSSIBLE FUNCTIONS
  ARE FERROXIDASE ACTIVITY, AMINE OXIDASE ACTIVITY, COPPER TRANSPORT
  AND HOMEOSTASIS, AND SUPEROXIDE DISMUTASE ACTIVITY.
CC -!- FUNCTION: MAY ALSO PLAY A ROLE IN FETAL LUNG DEVELOPMENT OR
  PULMONARY ANTIOXIDANT DEFENSE.
CC -!- CATALYTIC ACTIVITY: 4 Fe(2+) + 4 H(+) + O(2) = 4 Fe(3+) + 2 H(2)O.
CC -!- COFACTOR: BINDS 6 CU-IONS PER MOLECULE. THIS PROTEIN BELONGS TO
  THE MULTICOPPER OXIDASES WHICH CONTAIN THREE DISTINCT CU CENTERS
  KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED
  BINUCLEAR.
CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN LIVER AND SECRETED INTO THE
  PLASMA. ALSO CHOROID PLEXUS, YOLK SAC, PLACENTA, AND TESTIS; NOT
  IN STOMACH AND SMALL INTESTINE. FETAL LUNG AND LIVER.
CC -!- INDUCTION: BY INFLAMMATION.
CC -!- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF
  2 PLASTOCYANIN-LIKE REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L33869; AAA40917.1; -
DR EMBL; M80529; AAB65820.1; -
DR EMBL; J02670; AAA40914.1; ALT_SEQ.
DR EMBL; M14102; AAA40915.1; -
DR PIR; A35210; A35210.
DR PIR; A29564; A29564.
DR HSP; P00450; ILCW.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR002355; MultiCu-oxidase2.
DR Pfam; PF00394; Cu-oxidase; 3.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
DR Oxidoreductase; Copper; Metal-binding; Glycoprotein; Plasma; Repeat;
  Signal.
KW SIGNAL.
FT CHAIN 1 19 PROBABLE.
FT DOMAIN 20 1059 CERULOPLASMIN.
FT DOMAIN 20 356 F5/8 TYPE A 1.
FT DOMAIN 20 199 PLASTOCYANIN-LIKE 1.
FT DOMAIN 208 354 PLASTOCYANIN-LIKE 2.
FT DOMAIN 369 712 F5/8 TYPE A 2.
FT DOMAIN 564 710 PLASTOCYANIN-LIKE 3.
FT DOMAIN 724 1055 F5/8 TYPE A 3.
FT DOMAIN 724 894 PLASTOCYANIN-LIKE 4.
FT DOMAIN 902 1051 PLASTOCYANIN-LIKE 5.
FT DISULFID 173 199 BY SIMILARITY.
FT DISULFID 275 356 BY SIMILARITY.
FT DISULFID 528 554 BY SIMILARITY.
FT DISULFID 631 712 BY SIMILARITY.
FT DISULFID 868 894 BY SIMILARITY.
```

FT	METAL	120	120	COPPER (TYPE 2) (BY SIMILARITY).
FT	METAL	122	122	COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	179	179	COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	181	181	COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	988	988	COPPER (TYPE 1) (BY SIMILARITY).
FT	METAL	991	991	COPPER (TYPE 2) (BY SIMILARITY).
FT	METAL	993	993	COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	1033	1033	COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	1034	1034	COPPER (TYPE 1) (BY SIMILARITY).
FT	METAL	1035	1035	COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	1039	1039	COPPER (TYPE 1) (BY SIMILARITY).
FT	METAL	1044	1044	COPPER (TYPE 1) (BY SIMILARITY).
FT	CARBONYD	138	138	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	226	226	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	396	396	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	582	582	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	756	756	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	920	920	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	271	271	G -> A (IN REF. 2).
FT	CONFLICT	604	605	ED -> DN (IN REF. 2).
FT	CONFLICT	823	823	T -> S (IN REF. 2).
FT	CONFLICT	833	833	V -> L (IN REF. 2).
FT	CONFLICT	868	868	C -> V (IN REF. 2).
FT	CONFLICT	891	891	L -> R (IN REF. 2).
SQ	SEQUENCE	1059 AA,	120840 MM,	12BA3B990A0B95E3 CRC64;

Query Match 3.8%; Score 121.5; DB 1; Length 1059;
Best Local Similarity 20.6%; Pred. No.3.4;
Matches 114; Conservative 55; Mismatches 144; Indels 241; Gaps 34.

OY	10	FLGP--REEVSEPRCKLKSTESYVFHNHNSADFHRIQEKTDGNDWVPTIIDVGHSY	67
Db	92	FLGPYIKAEVGGKXSVHKNFASRPTTFHAH-----GVTY	126
OY	68	LOENK----ITTDLHRPLHDEMG-----NRDVVISDSQVLQEARPLVSAD	113
Db	127	TKANGAIPYDPTTFFORADDKLFGGOOYLVLRAINEPSRG-SNCTYRIHSHVDAP	185
OY	114	DEIVTSKAFTPIPKRPKKRNNGRNANHLNGINDGGQKEKFKNSEIDNELF	173
Db	186	KDIAS--GLISPL-----ILCKKSLNKK-----ENDIQEY	217
OY	174	QYKIEEL---LEKEKDFENSEKESEPSOEFPVEGHNNGLLRDEEKKDL--SN	226
Db	218	LMFSVDENLSWYLE--DNKTFCSEPP-----KYDKONEDQJESN	256
OY	227	K-----AMP--SHCD-----YOONLGNEPDKYCNGOVIFPCDTSFTSRPEK	268
Db	257	RMYISINGYTFGSLLPGISMCAEDRVKWLRLFGMGENVDVH-----	294
OY	269	OSVYPRTIYUGRPLSLANHLNIQRSGRPNSNIFOAO--DDSOI-ON-GYYVNC-	322
Db	295	SELF-----HGALTIYSKNHTDI-----INFRTLIDVSWAODPGVMWLSCON	339
OY	323	--HYVNMNCMTFOONNEYTCSESENRSVHPSGNCSMODRYVNSNGCFEVBRCWKDHCMKD	380
Db	340	LMHLAKGLDAFQ--VRQCNK-----PSPRD-DIQDHNVRIRUYIAAEETIM-DVA--	385
OY	381	HNGTDREFVNQOFQEELKLNKLKLLILRLPGSGKTL--SKILL--GONRDG----	430
Db	386	PAGTDTFTGTGFN-----TSLGSDSRKFVEFGATRIIGSGYKUL	422
OY	431	VSS--DDDFHHNODGYRYAVUNOLGAHDMMNONAKAALDOGRSPVLIIDNTNIDAMEBKPY	488
Db	423	VYREYTDSSF-----TNKRKGPRDEHLGI-----LGPV	451
OY	489	VEVALIGKGYVEFH	502
Db	452	IMAENVDIRIVTFH	465

RESULT 10
USOL_YEAST

ID	USO1.YEAST	STANDARD:	PRF:	1790 AA.
AC	P25386:			
DT	01-MAY-1992 (Rel. 22, Created)			
DT	01-MAY-1992 (Rel. 22, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Intracellular protein transport protein USO1.			
GN	USO1 OR INT1 OR YDL058W.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
OX	NCBI_TaxID:4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-X2180-1A:			
RX	MEDLINE=91185402; PubMed=2010462;			
RA	Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,			
RA	Yamasaki M.;			
RT	"A cytoskeleton-related gene, uso1, is required for intracellular			
RT	protein transport in Saccharomyces cerevisiae.";			
RL	J. Cell Biol. 113:245-260(1991).			
RN	[2]			
RP	SEQUENCE OF 782-1790 FROM N.A.			
RA	Hostetter M.R., Herman D.J., Bendel C.M., McClellan M., Tao N.,			
RA	Kendrick K.E.;			
RL	Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE OF 1-8 FROM N.A.			
RA	Bal Y., Symington L.S.;			
RL	Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI			
CC	COMPLEX.			
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR			
CC	MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE			
CC	ER AND THE GOLGI COMPLEX.			
CC	-1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED			
CC	OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL			
CC	COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.			
CC	-1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: X54378; CA38253.1;			
DR	EMBL: L03188; AAB00143.1;			
DR	EMBL: U53668; AAB66659.1;			
DR	PIR: A38455; A38455.			
DR	SGD: S0002216; USO1.			
DR	InterPro: IPR002017; Spectrin.			
KW	Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.			
FT	DOMAIN 1 724			
FT	DOMAIN 725 1790			
FT	DOMAIN 465 487			
FT	DOMAIN 991 1790			
FT	DOMAIN 1172 1786			
FT	CONFLICT 847 847			
FT	CONFLICT 924 924			
FT	CONFLICT 1253 1253			
FT	CONFLICT 1319 1319			
FT	CONFLICT 1461 1461			
FT	CONFLICT 1581 1581			
FT	CONFLICT 1600 1600			
FT	CONFLICT 1661 1661			
FT	CONFLICT 1772 1772			
SO	SEQUENCE 1790 AA; 206424 MW; 6CE3B216E9F5DA18 CRC64;			

```

Query Match      3.7%; Score 118.5; DB 1; Length 301;
Best Local Similarity 30.4%; Pred. No. 1.1;
Matches 46; Conservative 18; Mismatches 60; Indels 27; Gaps 7;

QY 402 KLILLRGLPGSGKTTLSRIILGQNRDGIVFSTDDY-----FHHQDGYRYNNQLGDAH 455
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2 KKIILITGCPGSGKSTWAREFIAKNPGCFYNINRDYRQSIMAHEERDEYKTKKKEGIVT 61
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 456 DWNONRAKQALDOGRS--PVIIDNTNIQ-----AWE--MKPYVEVAIGKGYRVEFHEPET 506
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 62 GMPEDTAKSILYIGDSVKYGVIIISDTNLNPPRLAWETFAKEY-----GKKVEHKYFEDV 114
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 507 WKWFDPELEKRN-KHGVSRKKIAQMLDRYE 536
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 115 PW-----TELVKRNSKRGTKAVPIDVLRSMYK 141
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
FRAL_MOUSE
ID_FRAL_MOUSE STANDARD: PRT: 273 AA

```

```

01-FEB-1996 (Rel. 33, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Pos-related antigen-1 (FRA-1).
FOSL1 OR FRA1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=BALE/CBYJ; TISSUE=Spleen;
RC MEDLINE=97047769; PubMed=8892610;
RX Huo L., Rothstein T.L.;
RA "Isolation and characterization of murine fra-1: induction mediated
RT by Cd40 and surface Ig is protein Kinase C dependent.";
RL J. Immunol. 157:3812-3818(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=97440303; PubMed=9294610;
RA Schreiber M., Poirier C., Franchi A., Kurzbaue R., Guenet J.L.,
RT Carle G.F., Wagner E.F.;
RA "Structure and chromosomal assignment of the mouse fra-1 gene, and its
RT exclusion as a candidate gene for oc (osteosclerosis).";
RL Oncogene 15:1171-1178(1997).
CC -!- SUBUNIT: HETERODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. FOS SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U34245; AAC52888.1; -
CC EMBL; AF017128; AAB71369.1; -.
CC HSSP; P01100; IFOS.
CC TRANSFAC; T00292; -.
CC MGD; MGI:107179; Fosl1.
CC InterPro; IPR000837; Leuzip_Fos.
CC InterPro; IPR004827; TF_BZIP.
CC Pfam; PF00170; bzip; 1.
CC PRINTS; PR00042; LEUZIPPRFOS.
CC SMART; SM00338; BRLZ; 1.
CC PROSITE; PS00036; BZIP_BASIC; 1.
CC Nuclear protein; DNA-binding.
CC DOMAIN 103 106
CC DNASIS 111 129
CC BASIC MOTIF.

```

```

FT DOMAIN 133 161 LEUCINE-ZIPPER.
FT DOMAIN 253 257 POLY-SER.
FT CONFLICT 17 17 P -> A (IN REF. 1).
FT CONFLICT 39 39 L -> F (IN REF. 1).
FT CONFLICT 170 170 P -> L (IN REF. 1).
FT CONFLICT 249 249 A -> T (IN REF. 1).
SQ SEQUENCE 273 AA; 29799 MW; 4B068B0259866805 CRC64;

Query Match 3.7%; Score 118; DB 1; Length 273;
Best Local Similarity 22.5%; Pred. No. 1;
Matches 60; Conservative 27; Mismatches 104; Indels 76; Gaps 11;

QY 87 PGNRPDVIESIDSOYLQAR-----PVLASADEI-YSTSKAFIGPI-----YKP-130
DB 17 PYGRAPPOQAQOQTAAQOKFHLVPSIDSSOELHMYQPHPLGPTGTPRLPAYIYQISPP 76
QY 131 -----PEKKRNEGRNEAHVINGIDRGQOKOKF-NSE 164
DB 77 QPRGVIALGPPGVRRRRRCQGISPEDEERRVRRENKLAACGRNRKELTDFLQAE 136
QY 165 KSEIDNELFQYKKEIELEKEKGFENSCKESEPOQFVPEYEGHNNGLKPDDEKKDL 224
DB 137 TDKLEDEKSGLOREIEELQOKERLE-----LYLEAHRPICKIPEGDKKDP 182
QY 225 SNKAMPSHCDYQONNGNEPDKPC-----NGQVI-----PFCDT-SFTSPFPMQSY 272
DB 183 GGGSGTSASSPPAPGR---PVPCISLSPGVLPEALHTPLMTPTSLPTPTPLVFTY 239
QY 273 PFIYVPGPLPSLNYHLNIQRFSGPPN 299
DB 240 P-----STPEPSSAHARRKSSSSGDDPS 261

RESULT 13
IF41_YEAST
ID IF41_YEAST STANDARD; PRT; 952 AA.
AC P39935.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Eukaryotic initiation factor 4F subunit p150 (eIF-4F) (mRNA cap-binding protein complex subunit p150).
GN TIF4631 OR YGRL62W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93330281; PubMed=8336723;
RA Goyer C., Altman M., Lee H.S., Blanc A., Deshmukh M., Woolford J.L., Trachsel H., Sonenberg N.;
RT "TIF4631 and TIF4632: two yeast genes encoding the high-molecular-weight subunits of the cap-binding protein complex (eukaryotic initiation factor 4F) contain an RNA recognition motif-like sequence and carry out an essential function.";
RL Mol. Cell. Biol. 13:4860-4874(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C;
RA MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae chromosome VII.";
RL Yeast 13:1077-1090(1997).
CC -I- FUNCTION: INTERACTS WITH THE MRNA CAP STRUCTURE, M7GPPX. THIS INTERACTION IS REQUIRED FOR EFFICIENT RIBOSOME BINDING TO THE MRNA. TIF4631 IS PROBABLY ESSENTIAL WHEN TIF4632 IS MISSING.
CC -I- SUBUNIT: THE CAP-BINDING PROTEIN COMPLEX IS COMPOSED OF AT LEAST TWO PROTEINS, A 24 KDA (TIF45) AND A 150-200 KDA SUBUNIT (TIF4631).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L16923; AAA02757.1;
DR EMBL: Z72847; CAA97184.1;
DR PIR: A48086; A48086.
DR SGD: S0003394; TIF4631.
DR InterPro: IPR003890; IF_eIF4G.
DR Pfam: PF02854; MIF4G; RNA_rec_mot.
DR SMART: SM00543; MIF4G; 1.
KW Initiation factor; Protein biosynthesis; RNA-binding;
KW Multigene family.
FT DOMAIN 173 200 PRO/SER-THR-RICH.
FT DOMAIN 375 384 ALA/GLU-RICH.
FT DOMAIN 488 553 ARG/SER-RICH.
FT DOMAIN 873 899 ARG/SER-RICH.
FT DOMAIN 869 872 POLY-GLU.
FT CONFLICT 7 7 H -> Q (IN REF. 1).
FT CONFLICT 37 37 T -> N (IN REF. 1).
FT CONFLICT 110 110 Q -> K (IN REF. 1).
FT CONFLICT 207 207 R -> K (IN REF. 1).
FT CONFLICT 361 361 D -> E (IN REF. 1).
SQ SEQUENCE 952 AA; 107101 MW; 391256802F86118E CRC64;

Query Match 3.7%; Score 117.5; DB 1; Length 952;
Best Local Similarity 19.5%; Pred. No. 5.3;
Matches 112; Conservative 85; Mismatches 217; Indels 159; Gaps 29;

QY 81 PLHDEPGRPDVIESIDSO-----VLQARPPVLSADE-IYSTKATIG 125
DB 423 PIEDVFSFNYEGIEGPDIKYKHEVKKYTGPTFLQFKDLNVKADAEWOSTASKIVI 482
QY 126 PIYKPEKKKKRNEGR-----NEAH-----VLNGIDRGQOKOKFENSEKSIDLELQF 175
DB 483 PEGMGGRNRSRDSGFGNNSSGHDPRNTSVANNMDRANSRSKRSRRMNDRRSNS 542
QY 176 YKEIELEKEDGFENSCKESE-----PSQEQFVPEYEGH-----NNGL 214
DB 543 Y--TSRRDRERSYRNEEKREDDKPEVAPLVPAANWVPRFKSKTEKKLAPDGKTEL 600
QY 215 LKPRD--EKKDLNRK-----AMPSHCDYQONLGNBPDKYPCNGQVIPTFCDTSTFSF 264
DB 601 LDKDEVERKMSLKLKLTLEMFDALISSEILAIANI-----SWETNGETLKAIVQIFLKA 656
QY 265 --RPEMOSYPTIYVYGPPLPSLNYHLNIQRFSGPPNPSNIFQADDSQIQNGIYVNNC 322
DB 657 CDEPHMSSYQAOLC--GKVVKELNPDIDETNEGTP-----KLVLYLVARC 703
QY 323 HVNMNMCTFDQNNETD--CSENRSYVPSNGSCSMODRYVNSGCEYERERCMRDHMD 379
DB 704 HAE-----FDKG--WTDKLPTEDEGTPLPE--IMSEYVAAASAKRGL----- 744
QY 380 KHNQDREYVNOQFOEKLKLD--KLI-----LRLGLPES--GKTLISRLILGQNDGIVF 432
DB 745 --GLVRFIGFLY--RLNLTGKMMFECEFRILMKDLTDSPEETLESVELLVNGEOLF 798
QY 433 STDQVFHNHDDYRVYVNOGDAHDNQNNAKQALIDGGRSPV-----IIDNTNIOAMKRP 487
DB 799 ETDLSF-----RTGGATLEGSOLLDSLGLILDNT--IQTAKISS 834
QY 488 YV-----EVALGKGYVEFHE--PETWAFDEPELEKRRKHGVS--KRIAMQMDR 534
DB 835 RIKFKLIDKELRHDKNNNSDKKDKGPKTIQIHHEEERQRLKNNNSNSNSRTNNSNR 894
QY 535 YEQY-----MSISIVANSVEPSHKSTQRPBP 562
DB 895 HSFRRDAPPAKSKDSFTTRTYISQNSQRAAPPK 927

```

RESULT 14

```
YF67_YEAST STANDARD; PRT; 1658 AA.
ID Q03661; Q04988;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 187.1 kDa protein in G0A1-ERG8 intergenic region.
GN YMR219W OR YMR261.13 OR YMR959.01.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RN SEQUENCE OF 1-711 FROM N.A.
RP STRAIN=S288C / AB972;
RC Dedman K., Brown D., Bowman S., Barrell B.G., Rajandream M.A.,
RA Walsh S.V.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 608-1648 FROM N.A.
RC STRAIN=S288C / AB972;
RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z49809; CAA89934.1; -
DR EMBL; Z49939; CAA90190.1; -
DR SGD; S0004832; YMR219W.
KW Hypothetical protein.
SQ SEQUENCE 1658 AA; 187137 MW; 3893F968305A757D CRC64;

Query Match 3.6%; Score 117; DB 1; Length 1658;
Best Local Similarity 18.5%; Pred. No. 12;
Matches 106; Conservative 84; Mismatches 176; Indels 206; Gaps 28;

QY 9 KFLGPREVTSEPRCKLSTTESYVFHNSNADFRIQKTDNDW-----VPVTVIDV 62
DB 322 KYMPRTDNTKIPVIEKYES-----DEHKVHORVSEDCAFDFGVSNIISVDDE 368

QY 63 RGHSLQENKIKTTDLRLPHDEMPCNRPDVIESDSQVLQEARPLIVSADDEIYSTKA 122
DB 369 SEDEESQAESYSANAENVYHHNEHELDKELTDIE-----SDSESQSAQES 416

QY 123 FIGP-----IVKPEKKRNEGRNAHVNLGINDRGQKEKQKFNSEKSIDNELFQYKE 178
DB 417 EQGSEDDFEYKMKNEKSTSE---ETENTSESDQGFAPKDAYTKNVEQGENDE----- 466

QY 179 IELEKEKOGFENSKESEPSOEQVFYFEGHNGLLKPDEEKKOLSNKAMPCHSDYQON 238
DB 467 -----EPEKDDIIRSLDKN-----FHGNNN---KSEISENVLENETDPAIVE-REN 509

QY 239 LGNEPKYPCNGQVITFCDTSTFRPEWQSVYPIVPGPLPSLNLHLNIQRFSGPP 298
DB 510 QINDVEGYDVGKSVES-----DLDEH----- 531

QY 299 NPPSNIFQAQDSQIOQNGYVNNCHVNNCMFTDONNEYTDCSENRSVHPS-----GNCC 354
DB 532 -SPDNLYDLAARMLO-----FQQRN-SNCPQKEEQVESYVLGHSNGS 573

QY 355 SMODRVVSGFCEVRRC-WKDHCMKHHG--TDR-----FVNQPFQEEKLN-KLOK 402
DB 574 NLSGRSLD-----ESEQIPLKDFGTGNNNNLKTDRGLDSSSVIEVEKYSKKLDGSTEK 629

QY 403 LLILLRGLPGSGKTTLSRILLQGNRQDGVFSTDDIFHHQDGYRYNNQLGDAHDNQNRA 462
```

```
Db 630 ELV-----PLSTDTTINSSLG--NEDSIYSLDD----- 657
QY 463 KQAIQGRSPVVIDN--TNIOAWEMK--PYVEVAIGKY-----RVFFHEP 504
Db 658 -----ADAISENLTDVPLMEIKTTPKVEVVISSEYSSYSTYEDNTVAMPPOVEYTS 709
QY 505 ETWKKFDP-----EELKRNKHGVSRRKKAQM 531
Db 710 ---FMNDPNSLNDYDEK--KHDLLKSTLAAL 736

RESULT 15
CN37_RAT STANDARD; PRT; 404 AA.
ID CN37_RAT
AC P13233;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 2',3'-cyclic nucleotide 3'-phosphodiesterase (EC 3.1.4.37) (CNP)
DE (CNPase).
OS Rattus norvegicus (Rat).
GN CNP.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87310616; PubMed=3040924;
RA Bernier L., Alvarez F., Norgard E.M., Raible D.W., Mentaberry A.,
RA Schember J.G., Sabatini D.D., Colman D.R.;
RT "Molecular cloning of a 2',3'-cyclic nucleotide 3'-phosphodiesterase:
RT mRNAs with different 5' ends encode the same set of proteins in
RT nervous and lymphoid tissues";
RL J. Neurosci. 7:2703-2710(1987).
CC -!- CATALYTIC ACTIVITY: Nucleoside 2',3'-cyclic phosphate + H(2)O =
CC nucleoside 2'-phosphate.
CC -!- SUBCELLULAR LOCATION: FIRMLY BOUND TO MEMBRANE STRUCTURES OF BRAIN
CC WHITE MATTER.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M18630; AAA40939.1; -
DR PIR; A45670; A45670.
DR InterPro; IPR001230; Prenyl_site.
KW Hydrolase; Membrane; Brain.
SQ SEQUENCE 404 AA; 45595 MW; 9C4CA888E6963BFF CRC64;

Query Match 3.6%; Score 116; DB 1; Length 404;
Best Local Similarity 24.7%; Pred. No. 2.2;
Matches 49; Conservative 32; Mismatches 65; Indels 52; Gaps 9;

QY 402 KLLILLRGLPGSGKTTLSRI-----LLQGNRQDGVFSTDDIFHHQDGYRYNV 448
DB 30 KTLFILRLGPGSGKSTLARNLNPWRSTTTAPRWCLLMLTRSFLA----- 72

QY 449 NOLGDAHDNQNRAKQAIQGRSP-----VIIDNTNIQAWEMKPYVEVAIGKYRVEPHE 503
DB 73 ---LGQTSPEYKRLDEDL-AGILPRDIRVLVDOTNHERERLDQFFEMADQYQVQLVE 129
QY 504 PETWKKFDEEELKRNKHGV-----SRKIAQMLDRYEQMSIS--IVMNSVEPSHKSTQ 556
DB 130 PKTANRLDCAQLKKNQWOLLARIDDLKLPKLEKDFLPYFGWFLTKKSETLRKDS- 188
QY 557 RPPPGQRQWCGSLGSH 574
DB 189 RP-----GSFQW--KLGNH 200
```


Tue Mar 11 10:11:24 2003

us-09-924-654-4.rsp

Page 11

Search completed: March 10, 2003, 17:53:28
Job time : 23 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 17:48:19 ; Search time 22 Seconds
(without alignments)
2547.562 Million cell updates/sec

Title: US-09-924-654-4

Perfect score: 3209

Sequence: 1 MSLGELEGKFLGPREVTSE.....RQRMGSLGSHNRVCVTNNH 563

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	195.5	6.1	232 2 F64576	hypothetical prote
2	164.5	5.1	162 2 T16163	hypothetical prote
3	161	5.0	149 2 D82178	hypothetical prote
4	154	4.8	400 1 ESBOP3	2',3'-cyclic-nucle
5	148.5	4.6	421 1 JCI317	2',3'-cyclic-nucle
6	147	4.6	420 1 ESM532	2',3'-cyclic-nucle
7	136.5	4.3	1551 2 F86342	F9H16.4 protein -
8	136	4.2	944 2 T18627	hypothetical prote
9	134.5	4.2	2269 2 A45555	hypothetical prote
10	134	4.2	1271 2 A45555	glutamate rich pro
11	133	4.1	420 2 I56377	2',3'-cyclic-nucle
12	130.5	4.1	2485 1 H71621	serine/threonine-s
13	128	4.0	964 2 T21865	hypothetical prote
14	128	4.0	2500 2 G71609	hypothetical prote
15	128	4.0	2924 2 T18378	variant-specific s
16	127	4.0	275 1 TVRTFR	transforming prote
17	127	4.0	1115 2 T41342	probable coiled-co
18	127	4.0	3119 2 T18414	protein g377 - mal
19	126	3.9	2269 2 T28677	rhopyri protein -
20	125	3.9	2523 2 T18477	hypothetical prote
21	123	3.8	1025 2 E86355	hypothetical prote
22	122.5	3.8	1804 2 H96597	hypothetical prote
23	121.5	3.8	2010 2 B71616	phosphatase (acid
24	120.5	3.8	871 2 D86355	protein T1615.12
25	120.5	3.8	1053 2 T51375	hypothetical prote
26	120.5	3.8	1059 1 A35210	ferroxidase (EC 1.
27	120	3.7	287 2 S45085	hypothetical prote
28	119	3.7	407 2 C45600	asparagine-rich bl
29	119	3.7	1790 2 S67593	transport protein

30	118.5	3.7	301 1 KIBPP4	polynucleotide kin
31	118.5	3.7	4981 2 T18489	hypothetical prote
32	118	3.7	969 2 T15446	hypothetical prote
33	117.5	3.7	952 2 S64473	translation initia
34	117.5	3.7	1844 2 D71612	hypothetical prote
35	117	3.6	773 2 F90537	lipoprotein limpor
36	117	3.6	1274 2 A89959	hypothetical prote
37	117	3.6	1658 2 S55101	hypothetical prote
38	116	3.6	404 2 A45670	2',3'-cyclic-nucle
39	115.5	3.6	1247 2 E71616	hypothetical prote
40	115	3.6	698 2 S49206	gl cyclin CLN1 - y
41	115	3.6	1164 2 S46769	hypothetical prote
42	115	3.6	1714 2 E71609	Ser/Thr protein ki
43	114.5	3.6	631 2 T32761	hypothetical prote
44	114	3.6	924 2 D81349	nitrate reductase
45	114	3.6	1044 2 S40704	hypothetical prote

ALIGNMENTS

RESULT 1
F64576
hypothetical protein HP0454 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 24-Nov-1999
C:Accession: F64576
R:Tomb, J.F.; White, O.; Kervavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalaf, H.G.; Glodek, A.; McKee
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney,
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: F64576
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-232 <TOM>
A:Cross-references: GB:AE000560; GB:AE000511; NID:92313554; PIDN:AAD07520.1; PID:9231
C:Genetics:
A:Start codon: TTG
C:Superfamily: Helicobacter pylori hypothetical protein HP0454

Query Match
Best Local Similarity 32.0%; Score 195.5; DB 2; Length 232;
Matches 57; Conservative 33; Mismatches 63; Indels 25; Gaps 8;

QY 394 EEKLNKLOKLLILRLGPGSGKTLRL-----LGNRDGIYFTDDYFHOD--CYR 445
DB 8 QRNRKSKKLVITINRAIPGGKSTLIQIEELAKSLGHSIS--VHSTDEYFIQDEGIR 65

QY 446 YNV---NOLGDAHDMNQNRKAKIDQGRPYIINTNTNIOAMENKPYEVAIGKRYVEFH 502
DB 66 HYVVDKKKLNLFYHQNQDEAFKQALENRDIYCCNTNENESQSKPYIDMAEFQYKILLI 125

QY 503 EPETWKMFDEPELEKRNKHG--VSR--KKIAQMDREYQNSISIVNNSVPSKHSIQ 566
DB 126 D-----FKNRHLETPMDYGMVDAQICKPRGIKHYDYELRVL--VEPODYERQ 175

RESULT 2
T16163
hypothetical protein F26A1.14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16163
R:Fulton, L.
submitted to the EMBL Data Library, May 1995
A:Description: The sequence of C. elegans cosmid F26A1.
A:Reference number: Z18469
A:Accession: T16163
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

```
A:Residues: 1-162 <FUL>
A:Cross-references: EMBL:U27312; NID:g860679; PID:g860692; PIDN:AAA68256.1; CESP:F26A1.1
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F26A1.14
A:Introns: 70/1; 125/2

Query Match      5.1%; Score 164.5; DB 2; Length 162;
Best Local Similarity 26.2%; Pred. No. 0.00071;
Matches 39; Conservative 35; Mismatches 46; Indels 29; Gaps 4;

Qy 394 BEKLNKLOKLL-----ILLRLPGSGKTTLSRLLGNRGGIVFSTDDYFHHQDGYRY 446
Db 13 DAEISDIQKCLVEGHTTILINGVTGSKSTLARELVHSENGVIVKND-----VSN 64

Qy 447 NVNQLGDAHDWNNQRAKQADQGRSPVIIDNTNIOAWEMKPYVEVAIGKGYRVEFHEPET 506
Db 65 NITR-----SVRFIDEDKHLIVVDENQVRSVKKFAELAVNGHVEIFVLEPDT 114

Qy 507 WVKFDPPELEKRNKH-----GYSRKKIAQM 531
Db 115 DWRHDAIECKRSEKDEIGSIESKIMOL 143

RESULT 3
D82178
hypothetical protein VC1610 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-feb-2001
C:Accession: D82178
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Hardison, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qln, H.; Dragol, I.; Sellers, B.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: R82035; MUID:20406833; PMID:10952301
A:Accession: D82178
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-149 <HEI>
A:Cross-references: GB:AB003852; NID:g9656130; PIDN:AAF94764.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1610
A:Map position: 1

Query Match      5.0%; Score 161; DB 2; Length 149;
Best Local Similarity 29.5%; Pred. No. 0.0011;
Matches 41; Conservative 26; Mismatches 46; Indels 26; Gaps 6;

Qy 404 LILLRLPGSGKTTLSRLLGNRGGIVFSTDDYFHHQDGYRYNNVNLGDAHDWNNQRA 462
Db 30 LTLIRLPGSGKSTLAKTL-----SAVHLEADMYFVNPQGEYHFRPEYLAQAHWCQQQT 84

Qy 463 KOAIDQGRSPVIIDNTNIOAWEMKPYVEVAIGKGYRVEF-----HEPETWKKFDPPELEK 517
Db 85 EYWLQGGKD-VVVSNTFVRHMEAVYRKLA--ROYRAKLTILVCRE-----RY 129

Qy 518 RNKHGVSRRKKIAQMLDRYE 536
Db 130 QNVHGVDEATVERMRQQWQ 148

RESULT 4
ESBOP3
2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) - bovine
N:Alternate names: CNPase; Cyclic-CMP phosphodiesterase; nucleoside-2',3'-cyclic-phospha
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 11-Jun-1999
C:Accession: A26861; A27038; A29136; A43792
R:Vogel, U.S.; Thompson, R.J.
Nucleic Acids Res 15, 7204, 1987
A:Title: Nucleotide sequence of bovine retina 2',3'-cyclic nucleotide 3'-phosphohydrolas
```

```
A:Reference number: A26861; MUID:88015580; PMID:2821502
A:Accession: A26861
A:Molecule type: mRNA
A:Residues: 1-400 <VOG>
A:Cross-references: GB:Y00405; NID:g261; PIDN:CAA68466.1; PID:g262
A:Experimental source: retina
R:Vogel, U.S.; Thompson, R.J.
FEBS Lett. 218, 261-265, 1987
A:Title: Molecular cloning of the myelin specific enzyme 2',3'-cyclic-nucleotide 3'-p
A:Reference number: A27038; MUID:87247281; PMID:3036592
A:Accession: A27038
A:Molecule type: mRNA
A:Residues: 1-400 <VO2>
A:Cross-references: GB:M27606; NID:g162880; PIDN:AAA30457.1; PID:g162881
A:Experimental source: brain
R:Kurihara, T.; Fowler, A.V.; Takahashi, Y.
J. Biol. Chem. 262, 3256-3261, 1987
A:Title: cDNA cloning and amino acid sequence of bovine brain 2',3'-cyclic-nucleotide
A:Reference number: A29136; MUID:87137604; PMID:3029107
A:Accession: A29136
A:Molecule type: mRNA
A:Residues: 1-317,'A',319-400 <KUR>
A:Cross-references: GB:J02659; NID:g162876; PIDN:AAA30456.1; PID:g162877
A:Note: this sequence has been corrected in reference A43792
R:Kurihara, T.; Fowler, A.V.; Takahashi, Y.
J. Biol. Chem. 262, 16754, 1987
A:Title: cDNA cloning and amino acid sequence of bovine brain 2',3'-cyclic-nucleotide
A:Reference number: A43792
A:Accession: A43792
A:Molecule type: mRNA
A:Residues: 316-320 <KU2>
A:Cross-references: GB:J02659
A:Note: this reference is a correction to reference A29136
C:Superfamily: 2',3'-cyclic-nucleotide 3'-phosphodiesterase
C:Keywords: phosphoric diester hydrolase

Query Match      4.8%; Score 154; DB 1; Length 400;
Best Local Similarity 28.6%; Pred. No. 0.012;
Matches 44; Conservative 37; Mismatches 49; Indels 24; Gaps 5;

Qy 394 BEKLNKLO--KLLILLRLPGSGKTTLSRLLGNRGGIVFSTDDYFHHQDGYRYNNVNL 451
Db 20 EETVATLQECTLFLIRGLPGSGKSTLAFIVDKYRDGCKMVSADSYKITPGARGSFSE- 78

Qy 452 GDAHDWNNQRAKQADQGRSP-----VIIDNTNIOAWEMKPYVEVAIGKGYRVEFHE 503
Db 79 ----EYKO-----LDEDLAACRRDRFVLYLDDTNHERERLEQLFELADQYQVVLVE 128

Qy 504 PETWKKFDPPELEKRNKHGVS--KKIAQMLDR 534
Db 129 PKTAWRLDCAQLKEKNQWLSADDLKLRPGLEK 162

RESULT 5
JC1517
2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) long splice form - human
N:Alternate names: 2',3'-cyclic nucleotide-3'-phosphohydrolase; CNPase; cyclic-CMP ph
N:Contains: 2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) short splice f
C:Species: Homo sapiens (man)
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 19-Jan-2001
R:Monoh, K.; Kurihara, T.; Takahashi, Y.; Ichikawa, T.; Kumanishi, T.; Hayashi, S.; M
Gene 129, 297-301, 1993
A:Title: Structure, expression and chromosomal localization of the gene encoding huma
A:Reference number: JC1517; MUID:93314977; PMID:8392017
A:Accession: JC1517
A:Molecule type: DNA
A:Residues: 1-421 <MON>
A:Cross-references: DDBJ:D13144; DDBJ:D13145; DDBJ:D13146; NID:g219398; PIDN:BAA39694
A:Note: long form (isoform 1)
A:Accession: JC1518
A:Molecule type: DNA
A:Residues: 21-421 <MO2>
```


QY 61 DVGRHSYLO--ENKIKTTDLHRPLHDEMPGNRPDVI-----ESTD-SQVLOEARPPPLY 110
Db 1208 ENNGKEVAKPTOKSOTTTSKRAVPDQP---PSIVTELVSKEEIEKATPEEPEPPKL 1264
QY 111 SADDEIYSTSKAFIGIYKPPKRRNEG--RNEAHLVINGINDRGQKQKFNSEKSEI 168
Db 1265 TKEEBEL-----IKKEEKKRQKAAKMQHRLBEIAKAEAMERKKKREKAKA 1315
QY 169 DNELFOFYKEIELEKEKGFNSC-----KESEPQOFVPPFYEGHNGL 214
Db 1316 -RAVLKAQAEAREKVKAYFCHSVIQIGFRSVSLARTREEAKEEKGKGFYIRDSN- 1373
QY 215 LKPDEKDLNSKAMPSCDYNLGNEDPKYPCNGOVIPTFCDDTSET 262
Db 1374 RPNDSRNCNRPNKGRNSKTRNGE-----STNEEISQTFITVSET 1417
RESULT 8
T18627
hypothetical protein B0001.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18627
R:Sims, M.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z18999
A:Accession: T18627
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-944 <WIL>
A:Cross-References: EMBL:Z69634; PIDN:CAA93451.1; GSPDB:GN00022; CESP:B0001.2
A:Experimental source: clone B0001
C:Genetics:
A:Gene: CESP:B0001.2
A:Map position: 4
A:Introns: 18/2; 75/3; 119/2; 350/3; 475/2; 487/3; 521/2; 601/3; 641/3; 749/2; 870/1

Query Match 4.2%; Score 136; DB 2; Length 944;
Best Local Similarity 20.7%; Pred. No. 0.56;
Matches 131; Conservative 83; Mismatches 205; Indels 214; Gaps 33;
QY 1 MSYGEIE-----GKFLGPREEVTSEPRCKKLKSTTE--SYVF-----HNHSNADF--- 43
Db 137 LSDGKIEVISKTAGVLTETKRNKSNKSEAKSEKYEYKIIIGVVTGKNORNAVFTCN 196
QY 44 -----HRIQKGTG---NDWPVPTIIDVRGHSYLOENKIKTTDLHRPLHDEMPGNRPDI 94
Db 197 LTVPGTDGIIKSQLQIGDWVEIRTKDFVKYFSPNPTADTPRENVKHFKE-----I 249
QY 95 ESIDSOVLOEARPLVSADDEIYSTSKAFIGIYKPPKRRNEGNEAHLV-NGINDR- 152
Db 250 EKIEH-----YSVTWAGN-----POVKIRNFOPFAGHVQGSVEDRF 287
QY 153 -GGQKEKQKFNSEKSEID-----NELFO-----FYKETELEKEKDGFNSECKESE 197
Db 288 LGTIAISKIVTSEKVDVLIKRLFOAGDGKQATWTFKKNVKEKSESDSGDSENNENS 347
QY 198 PSQOEFVPPFYEGHNGNLLKPDDEKDLNSKAMPSCDYNLGNEDPKYPCNGOVIPTFC 257
Db 348 SSTDS-----RPSRPNKVENKKS-----EEKINNSSEK----- 376
QY 258 DTSFTSFRPEWQSVFPFIVYGPPLPS-LNYHLNLTQRFSGPPN-----PPSIFQAQDSD 311
Db 377 -----KEYSNVPKINSHLPPEPSPMNQ-----GPPGIMHPRPIGHGAIPNP 422
QY 312 QIQNGYYVNNCHVNNCMFTFDONNEYTDCSENRSVHPSGNCSQMDRVVSGFCFVRER 371
Db 423 LMYSPYMSMAIPQOM-----NRQVPTNNSPEVHIHKG----- 458
QY 372 CWKDHCMKHNKGTDRF-----VNOQFOEELKLNKLOKLLIL-LRGLPGSGKTTLSRILLGN 426
Db 459 -FKP-LISTENPLVFRGRCMCFSPESQKTNVLRKAVITSLK--PNNKGH-----LYGKE 509

QY 427 RDGIVFSTDDYFHHODGYRYNVNQ-LGDAHDWNNQRAKQAIQDGRSPVLIIDNTNQAMEM 485
Db 510 KTAFLWLDDD--HKOSVYVYVSKNDGIEPGHPFNGLFA-----SNGDKWEC 552
QY 486 KPYVEVAIGK-----GYRVEFHEPETWKFDPPELEKRNK- 520
Db 553 KKYVK-PLGKLMGIVSNSDIELQLIVETYPQSGERL---NPETYHSFIGIVDKFNKL 608
QY 521 -----HGVSRK-KIAQMLDRYQMSISIVMS 547
Db 609 PEDCSRGVKYKISIKMWNIVERNEWCWIVSKVFS 641
RESULT 9
T18472
hypothetical protein C0440c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T18472
R:Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z18937
A:Accession: T18472
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-2269 <LAW>
A:Cross-References: EMBL:AL008970; NID:e1407852; PID:e1332566; PIDN:CAA15615.1
C:Genetics:
A:Gene: C0440c
A:Map position: 3
Query Match 4.2%; Score 134.5; DB 2; Length 2269;
Best Local Similarity 19.3%; Pred. No. 2.3;
Matches 124; Conservative 82; Mismatches 210; Indels 225; Gaps 30;
QY 24 KKLKSTTESYFHNHNSADFHRIQE-KTGNWVPVTIIDVRGHSYLOENKIKTTDLHRPL 82
Db 9 KRRRTYVESYNIINIAKRYCMCNKMDND---NNIINEKKY-FNNSSSIKNT----- 58
QY 83 HDEMPGNRPDIESTDSQVLOEARPLVSADDEIYSTSKAFIGIYKPPKRRNEGNE 142
Db 59 -----GNNYK-NDNIDKSYDMCERSMMDROENIYNN-----YK---NKNRNNYSNK 102
QY 143 AHVLNGIN-----DRGGQKE-----KOKF----- 161
Db 103 NYSNNIINHMNDNRRTYKGDLDGRKQYLSNKHYESKNDNENSYNLKYNFNIPPIKYNN 162
QY 162 -----NSEKSEIDNELFOFYKEIELEKEK-----DGFENSC-----KESEPQEQ 202
Db 163 NNNNNNNNNORQNDNNIYKFYDDKNEKKSNTKTHSYGNFNNFNTNDPKQNTMSNES 222
QY 203 FVPFVEGH-----NNGLLKPDDEKDLNSKAMPSCDYNLGNEDPKYPCNGQVI 253
Db 223 HMTODGHPYIYNNHNNNDGRK---NRKDISLSYGYVYKISNDNLSKENTSY----- 272
QY 254 PTFCDTSFTSFRPEWQSVFPFIVYGPPLPSLNYHLNLTQRFSGPPNPSNIFQAODSQI 313
Db 273 -----INQYNN-----SGGAN--RHVINAQHNQ 295
QY 314 QNGYYVNNCHVNNCMFTFDONNEYTDCSENRSVHPSGNCSQMDRVVSGFCFVRER 373
Db 286 HNNNOHNN---NFNYHNYKNAKPYQDKYKNGDSFKFSAYNABQHDHNMKNYQDTHKNNM 352
QY 374 KDHCMD-----KHNGT---DRFVNOQFOEELKLNKLOKLLIL-LRGLPGSGKTTLSRILLGN 426
Db 353 KHSKGTDTAIFNGOKVFGSKNKKQKERR-KPEK-----NKERGSSI-----N 395
QY 427 RDGIVFSTDDYFHHODG---YRYNVNQLGDA-----HDWNNQRAKQAIQD-GRSPVII 475
Db 396 RPYHNNNNNNNNNNNNNNYNNHNFCDDEARRNNKNNENWCEINSNELYKDGDMPIYN 455
QY 476 DNTNLQAMEMKPYVEVAIGKGYRVEFHEPETWKFDPPELEK-----RNKHGV----- 523

C:Accession: T18378
R:Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bl, X.; Ma, X.C.; Feldman, M.; Taraschi, T.
Cell 82, 77-87, 1995
A:Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and
A:Reference number: Z18925; MUID:95330812; PMID:7541722
A:Accession: T18378
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2924 <BAR>
A:Cross-references: EMBL:U27338; NID:g914918; PID:g914919; PIDN:AAB60251.1
C:Genetics:
A:Gene: EMP1
A:introns: 2476/3

Query Match 4.0%; Score 128; DB 2; Length 2924;
Best Local Similarity 20.8%; Pred. No. 8.5;
Matches 88; Conservative 56; Mismatches 153; Indels 126; Gaps 22;

```
QY 117 YSTSKAFIGFIYKPEKKRNEGNEAHVNLGIN----- 150
      | | | | | | | | | | | | | | | | | | | |
Db 1472 YNLIKHKRIDPCI-----KKKQDKTEHKCINGCNICEVYRKMLEIKNGEMGNIKKHYN 1526
      | | | | | | | | | | | | | | | | | | | |
QY 151 DRGGOKKQKFNSEKSEIDNELFOF-YKEIEE-LEKEKGFENSCKESEPSEQFVPFYE 208
      | | | | | | | | | | | | | | | | | | | |
Db 1527 NSNDDEKTIYVYKSYFVDGLFDTDYKKAQKVEDEK-----ERKKWGCT 1573
      | | | | | | | | | | | | | | | | | | | |
QY 209 GHNNGLKPDPEKKDLSENKAMPSSHCDYQONLGNPEPKYPCNGQVPTFCDTSTSPREP 268
      | | | | | | | | | | | | | | | | | | | |
Db 1574 GHDCSEKEKEKNFTNLI---SELQDKITSCQNKHNPNKG---TACD-PFSPPTPE- 1625
      | | | | | | | | | | | | | | | | | | | |
QY 269 QSYVPETVPYGPPLPSLNYHLNIORESGPPNP-----SNTFOAQDDSOIONGYVNN 321
      | | | | | | | | | | | | | | | | | | | |
Db 1626 -ETDPLDDPTDPLDD-DQHTEQPKFCPPPPPTCYEKIAKELRAEGRKINNELKNG 1683
      | | | | | | | | | | | | | | | | | | | |
QY 322 CHVMNMGTFDONNEYTDCSENSSVHPSGNGCSMODRYVNGFCVEYERQWKDCHMDKH 381
      | | | | | | | | | | | | | | | | | | | |
Db 1684 KDEGKCNNAVKKNGAVIGES-----CKFEQTY-ENSVNINNNKC-----KD 1725
      | | | | | | | | | | | | | | | | | | | |
QY 382 NGTDRF-VNOQFOEEKLNKLQKLIL-----LRGLPGSGKTTLS----- 419
      | | | | | | | | | | | | | | | | | | | |
Db 1726 NQNERFKIGKWNKYGITIRKDLCLIPRREHMLDLSMLGRTTISDSALLKKIQEAA 1785
      | | | | | | | | | | | | | | | | | | | |
QY 420 -----RILGQNRDGIYFSTDDYFHQ--DGIYINVNQGD---AHD-WNQNRAKQA 465
      | | | | | | | | | | | | | | | | | | | |
Db 1786 KSERDDIIRKLLLEQN-----SCDE---HRICDAMKYSFADLGDIIIRGDLNKNKQKG 1836
      | | | | | | | | | | | | | | | | | | | |
QY 466 IDQ 468
      | | | | | | | | | | | | | | | | | | | |
Db 1837 LQK 1839
```

Search completed: March 10, 2003, 17:54:55
Job time : 29 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 17:54:19 ; Search time 20 Seconds
(without alignments)
1229.260 Million cell updates/sec

Title: US-09-924-654-4
Perfect score: 3209
Sequence: 1 MSYGEIEGKFLGPREVTSF.....RQWGGSLGSHNVCVTNNH 583

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188354 seqs, 42170167 residues
Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3209	100.0	583	10	US-09-924-654-4
2	3205	99.9	594	10	US-09-925-900-1079
3	449.5	14.0	132	10	US-09-924-654-10
4	372	11.6	104	10	US-09-924-654-12
5	120	3.7	1111	10	US-09-815-242-12955
6	109.5	3.4	1338	10	US-09-402-100-4
7	108.5	3.4	291	10	US-09-925-900-968
8	108	3.4	665	9	US-09-820-843A-107
9	106	3.3	506	10	US-09-976-165-19
10	105	3.3	1093	10	US-09-801-968-392
11	104.5	3.3	2125	10	US-09-919-172-29
12	104	3.2	1435	9	US-10-153-273-4
13	104	3.2	1863	9	US-10-022-819-2
14	104	3.2	4636	10	US-09-835-996A-33
15	103	3.2	819	10	US-09-825-144-14
16	103	3.2	861	9	US-09-820-843A-109
17	102.5	3.2	1139	9	US-09-820-843A-15
18	102.5	3.2	1421	10	US-09-924-154-13
19	101	3.1	2507	9	US-09-819-104A-2

20	100	3.1	1295	10	US-09-726-949A-1	Sequence 1, Appl1
21	100	3.1	3594	9	US-10-150-821-4	Sequence 4, Appl1
22	100	3.1	3594	10	US-09-911-842-4	Sequence 4, Appl1
23	99.5	3.1	807	9	US-10-029-217A-2	Sequence 2, Appl1
24	99.5	3.1	935	9	US-10-029-217A-31	Sequence 31, Appl1
25	99.5	3.1	935	9	US-10-029-217A-32	Sequence 32, Appl1
26	99	3.1	340	9	US-09-789-054A-6	Sequence 6, Appl1
27	98.5	3.1	1719	9	US-10-024-450-4	Sequence 4, Appl1
28	98.5	3.1	3571	9	US-10-150-821-2	Sequence 2, Appl1
29	98.5	3.1	3571	10	US-09-911-842-2	Sequence 2, Appl1
30	98	3.1	532	10	US-09-891-160-2	Sequence 2, Appl1
31	98	3.1	782	9	US-09-935-868-48	Sequence 48, Appl1
32	98	3.1	782	9	US-09-935-868-52	Sequence 48, Appl1
33	97.5	3.0	1805	9	US-09-820-843A-73	Sequence 73, Appl1
34	97	3.0	370	9	US-10-086-823-8	Sequence 8, Appl1
35	97	3.0	370	9	US-10-139-583-37	Sequence 37, Appl1
36	97	3.0	370	9	US-10-039-847A-2	Sequence 2, Appl1
37	97	3.0	370	9	US-09-823-033-5	Sequence 5, Appl1
38	97	3.0	370	10	US-09-808-972-2	Sequence 2, Appl1
39	97	3.0	370	10	US-09-915-582-56	Sequence 56, Appl1
40	97	3.0	629	10	US-09-799-875-17	Sequence 17, Appl1
41	97	3.0	843	10	US-09-815-242-12552	Sequence 5370, Ap
42	97	3.0	843	10	US-09-815-242-12552	Sequence 12552, A
43	97	3.0	843	10	US-09-815-242-12898	Sequence 12898, A
44	97	3.0	1086	10	US-09-924-154-15	Sequence 15, Appl1
45	97	3.0	1863	9	US-09-734-672-2	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-09-924-654-4
: Sequence 4, Application US/09924654
: Patent No. US20020146712A1
: GENERAL INFORMATION:
: APPLICANT: Goold, Richard D.
: APPLICANT: Aketdiom, Ingrid E.
: APPLICANT: Sellhammer, Jeffrey J.
: APPLICANT: Coleman, Roger
: TITLE OF INVENTION: TUMOR SUPPRESSOR
: FILE REFERENCE: PC-0049 CIP
: CURRENT APPLICATION NUMBER: US/09/924, 654
: CURRENT FILING DATE: 2001-08-07
: NUMBER OF SEQ ID NOS: 21
: SOFTWARE: PERL Program
: SEQ ID NO 4
: LENGTH: 583
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No. US20020146712A1 496071CD1
US-09-924-654-4

Query Match      100.0%      Score 3209; DB 10; Length 583;
Best Local Similarity 100.0%; Pred. No. 9.7e-247;
Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSYGEIEGKFLGPREVTSFRRCKKLTSTSYFHNHNSNADFRIOKGTGNDWPVTII 60
DB 1 MSYGEIEGKFLGPREVTSFRRCKKLTSTSYFHNHNSNADFRIOKGTGNDWPVTII 60
QY 61 DVRGHSYLQENKIKTTDLHRPLHDEMPGNRPDVIESIDSOVLQEARPLVSADEIVSTS 120
DB 61 DVRGHSYLQENKIKTTDLHRPLHDEMPGNRPDVIESIDSOVLQEARPLVSADEIVSTS 120
QY 121 KAFIGPIYKPEPKKRNKNGRNEAHVNLGINDRGCKEKQKNSKSELDNLFQFYKEIE 180
DB 121 KAFIGPIYKPEPKKRNKNGRNEAHVNLGINDRGCKEKQKNSKSELDNLFQFYKEIE 180
QY 181 ELEKXDFENSKSESESOEFPVFEYGHNNGLKPKPEKKDLSNKKAMPCHCYOOWLG 240
DB 181 ELEKXDFENSKSESESOEFPVFEYGHNNGLKPKPEKKDLSNKKAMPCHCYOOWLG 240
```

```
Db 181 ELEKEKGFENSCKESEPSQEQFVPFYECHNGLLKPDDEKKDLNKAHPSHCDYQOQNLG 240
QY 241 NEPDYKPCNGQVIFPCDTSTSFPEWQSVYFFIVPGPLPSLNYHLNIQRFSGPPNP 300
Db 241 NEPDYKPCNGQVIFPCDTSTSFPEWQSVYFFIVPGPLPSLNYHLNIQRFSGPPNP 300
QY 301 PSNIFQAQDDSOIQNGYVYNNCHVNNCMTFDQNNNEYTDCSENRSVHPSPGNGCSWQDRY 360
Db 301 PSNIFQAQDDSOIQNGYVYNNCHVNNCMTFDQNNNEYTDCSENRSVHPSPGNGCSWQDRY 360
QY 361 VSGFCEVRERCWKDCHMDKHNGTDREVNQOQFQEBKLNKQLKLLILRLGLPGSGKTTLSR 420
Db 361 VSGFCEVRERCWKDCHMDKHNGTDREVNQOQFQEBKLNKQLKLLILRLGLPGSGKTTLSR 420
QY 421 ILLGNRQDGIIVSTDYFHHQDGYRYNNQLGDAHDWNONRAKQAIDQGRSPVIIDNTNI 480
Db 421 ILLGNRQDGIIVSTDYFHHQDGYRYNNQLGDAHDWNONRAKQAIDQGRSPVIIDNTNI 480
QY 481 QAWEMKPYVEVAIGKGYRVEFHEPETWKFDPPELEKRNKHGVSRRKKAQMLDRVEYQMS 540
Db 481 QAWEMKPYVEVAIGKGYRVEFHEPETWKFDPPELEKRNKHGVSRRKKAQMLDRVEYQMS 540
QY 541 ISIVNSVEPSHKSTQRPPOGQRWGSGSLGSHNRVCVTNNH 583
Db 541 ISIVNSVEPSHKSTQRPPOGQRWGSGSLGSHNRVCVTNNH 583

RESULT 2
US-09-925-300-1079
; Sequence 1079, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCR/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1079
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (430)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1079

Query Match 99.9%; Score 3205; DB 10; Length 594;
Best Local Similarity 99.8%; Pred. No. 2.1e-246;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSYGEIEGKFLGPREEVTSRPRCKKLKSTTSYVFHNHNSADFHRIQKGTGNDWVPVTII 60
Db 12 MSYGEIEGKFLGPREEVTSRPRCKKLKSTTSYVFHNHNSADFHRIQKGTGNDWVPVTII 71

QY 61 DVRGHSYLOENKIKTTDLHRPLHDEMPGNRPDVIETSDSQVLQEARPLVSADDEIYSTS 120
Db 72 DVRGHSYLOENKIKTTDLHRPLHDEMPGNRPDVIETSDSQVLQEARPLVSADDEIYSTS 131

QY 121 KAFIGIYKPKPKKRNKGRNEAHLNGINDRGQKQKQKENSEKSETDNELFQFYKIE 180
Db 132 KAFIGIYKPKPKKRNKGRNEAHLNGINDRGQKQKQKENSEKSETDNELFQFYKIE 191

QY 181 ELEKEKGFENSCKESEPSQEQFVPFYECHNGLLKPDDEKKDLNKAHPSHCDYQOQNLG 240
Db 192 ELEKEKGFENSCKESEPSQEQFVPFYECHNGLLKPDDEKKDLNKAHPSHCDYQOQNLG 251
```

```
QY 241 NEPDYKPCNGQVIFPCDTSTSFPEWQSVYFFIVPGPLPSLNYHLNIQRFSGPPNP 300
Db 252 NEPDYKPCNGQVIFPCDTSTSFPEWQSVYFFIVPGPLPSLNYHLNIQRFSGPPNP 311
QY 301 PSNIFQAQDDSOIQNGYVYNNCHVNNCMTFDQNNNEYTDCSENRSVHPSPGNGCSWQDRY 360
Db 312 PSNIFQAQDDSOIQNGYVYNNCHVNNCMTFDQNNNEYTDCSENRSVHPSPGNGCSWQDRY 371
QY 361 VSGFCEVRERCWKDCHMDKHNGTDREVNQOQFQEBKLNKQLKLLILRLGLPGSGKTTLSR 420
Db 372 VSGFCEVRERCWKDCHMDKHNGTDREVNQOQFQEBKLNKQLKLLILRLGLPGSGKTTLSR 431
QY 421 ILLGNRQDGIIVSTDYFHHQDGYRYNNQLGDAHDWNONRAKQAIDQGRSPVIIDNTNI 480
Db 432 ILLGNRQDGIIVSTDYFHHQDGYRYNNQLGDAHDWNONRAKQAIDQGRSPVIIDNTNI 491
QY 481 QAWEMKPYVEVAIGKGYRVEFHEPETWKFDPPELEKRNKHGVSRRKKAQMLDRVEYQMS 540
Db 492 QAWEMKPYVEVAIGKGYRVEFHEPETWKFDPPELEKRNKHGVSRRKKAQMLDRVEYQMS 551
QY 541 ISIVNSVEPSHKSTQRPPOGQRWGSGSLGSHNRVCVTNNH 583
Db 552 ISIVNSVEPSHKSTQRPPOGQRWGSGSLGSHNRVCVTNNH 594

RESULT 3
US-09-924-654-10
; Sequence 10, Application US/09924654
; Patent No. US20020146712A1
; GENERAL INFORMATION:
; APPLICANT: Goold, Richard D.
; APPLICANT: Akerbloom, Ingrid E.
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: TUMOR SUPPRESSOR
; FILE REFERENCE: PC-0049 CIP
; CURRENT APPLICATION NUMBER: US/09/924,654
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020146712A1 g533948
US-09-924-654-10

Query Match 14.0%; Score 449.5; DB 10; Length 132;
Best Local Similarity 82.1%; Pred. No. 1.4e-28;
Matches 87; Conservative 0; Mismatches 0; Indels 19; Gaps 1;

QY 1 MSYGEIEGKFLGPREEVTSRPRCKKLKSTTSYVFHNHNSADFHRIQKGTGNDWVPVTII 60
Db 46 MSYGEIEGKFLGPREEVTSRPRCKKLKSTTSYVFHNHNSADFHRIQKGTGNDWVPVTII 86

QY 61 DVRGHSYLOENKIKTTDLHRPLHDEMPGNRPDVIETSDSQVLQEAR 106
Db 87 DVRGHSYLOENKIKTTDLHRPLHDEMPGNRPDVIETSDSQVLQEAR 132

RESULT 4
US-09-924-654-12
; Sequence 12, Application US/09924654
; Patent No. US20020146712A1
; GENERAL INFORMATION:
; APPLICANT: Goold, Richard D.
; APPLICANT: Akerbloom, Ingrid E.
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: TUMOR SUPPRESSOR
```

```
FILE REFERENCE: PC-0049 CIP
CURRENT APPLICATION NUMBER: US/09/924,654
CURRENT FILING DATE: 2001-08-07
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PERL Program
SEQ ID NO 12
LENGTH: 104
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020146712A1 PANC 1B
NAME/KEY: unsure
LOCATION: 5,44, 75, 101
OTHER INFORMATION: unknown or other
US-09-924-654-12

Query Match
Best Local Similarity 94.6%; Pred No. 1.4e-22; Length 104;
Matches 70; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSYGEIEGKFLGPREVTESEPRCKKLKSTESYVFHNHNSADFHRIQKGTGNDVPEVTII 60
DB 31 MSYGEIEGKFLGPREVTESEPRCKKLKSTESYVFHNHNSADFHRIQKGTGNDVPEVTII 90
QY 61 DVRGHSYLOENKIK 74
DB 91 DVRGHSYLOENKIK 104

RESULT 5
US-09-815-242-12955
Sequence 12955, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12955
LENGTH: 1111
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12955

Query Match
Best Local Similarity 18.1%; Pred No. 0.38; Length 1111;
Matches 117; Conservative 69; Mismatches 224; Indels 238; Gaps 25;
```

```
QY 3 YGEIEGKFLGPR-----EVTSEPRCKKLKSTESYVFH-----NHSNADFHRIQKGTGN 52
DB 47 YSRPRGKFRPMASVAYENENEQADTISDEKQY--HRDYKQSHDSRSQKRRRRRND 104
QY 53 DWVPVTTIDVRGHSYLOENKIKTTDLRPLHDEMPGNRPDIRESISOVLQ-----103
DB 105 TTEEQNYSEQRGNSKISQOSIKYKD-HSHYHTKPKPTIYSAINGIKETHKPTTHMYSN 163
QY 104 ---EAPPLVSADDEIYSTSK---AFIG-----PIYKPEKKRRNEGNEAH 144
DB 164 NTNHRAKDSTPDYHKSEFSTSEVPSAIFGTMKPKLNGRIPIVSKSEKVESDKQYDX 223
QY 145 VLVGINDRGQKQKQKFN-----EKSEIDNEL-----F 173
DB 224 VAKTQTSQNKQLEQEKQNDQSVYKQGTASKSSDQENVSSTKSMQNYKVDNTIKIENIYAS 283
QY 174 QFYKET-----EELKEKQDFGNSCKSESPSO-EOFPVPEYG 209
DB 284 QIYEELRREPERKVLQKRFKALQOKREHKNEQDQAIQRAIDEMAKQAEKYV----- 338
QY 210 HNNGLKLPDEKKDLNKAHPSHCDYQON--LGNEPDKYPC--NGOVIPTFCDTSFTSPR 265
DB 339 -GDSLSLNDSDLDLNDSTASQLHTNGIENETYSNDENKQASIQNEDTNDTHDES----- 392
QY 266 PEMQSVYPTIVPGPLPSLNLHLNIOFSGPPNPSPNIFQADDSQIQNGIYVNNCHVN 325
DB 393 -----PYNEEVSINQVSTTKQSLDDEVTVSNV-TSQHQSALQHNVEVN----- 435
QY 326 WNCMTFDQNNETDCESENSSVHPSGNGCSMDQRYVSNQCEVREKQWCHDKHNGTD 385
DB 436 -----DK-----D 438
QY 386 RFVNOQFOEEKLNKLQKLLILLGLPGSGKTLRIILGQNRDGIYSTDDYFHHDGYR 445
DB 439 ELKNO-----SRLIADSEEGATNKEE-----460
QY 446 YVNVQLGDAHDMNQNRKAQIDGSRPVIIDNTNIGAMENKPYEVAIGKYVEFHEPE 505
DB 461 YSGSQIDDAEFYELNDTE--VDEDITSNIEDNTNRNASEM--HVDAPKQGEYAVTESQVN 516
QY 506 TMMKFPDELEKRNKHGVSRRKIAQMLDRYEYMSI-SIVMNVSEESH 552
DB 517 NDKTVDNIEIELAPRH---KKDDQ-----NLVSNLSKTDVNDNH 554

RESULT 6
US-09-402-100-4
Sequence 4, Application US/09402100
Patent No. US20010019834A1
GENERAL INFORMATION:
APPLICANT: Daewoong Pharmaceutical Co, LTD
APPLICANT: Kim, Byung-O
APPLICANT: Shin, Sung-Seup
APPLICANT: Yu, Young-Hyo
APPLICANT: Park, Myung-Hwan
APPLICANT: Choi, Deok-Joon
APPLICANT: Jung, Hyung-Jin
TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Hel
FILE REFERENCE: 0136/00140
CURRENT APPLICATION NUMBER: US/09/402,100
CURRENT FILING DATE: 1999-09-27
EARLIER APPLICATION NUMBER: KR 97-11950
EARLIER FILING DATE: 1997-03-31
EARLIER APPLICATION NUMBER: KR 97-11951
EARLIER FILING DATE: 1997-03-31
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 1338
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
```

; OTHER INFORMATION: CagA/CTXA2B Chimeric protein
US-09-402-100-4

Query Match	3.4%	Score 109.5	DB 10	Length 1338
Best Local Similarity	19.1%	Pred. No. 3.4		
Matches 128	Conservative	98	Mismatches 240	Indels 205
			Gaps 33	

```
QY      3 YGIECKFLGPREEVTSRPRCKKLKSTTSEYVFHNHNSNADFRIQEKTGNDWVPVTIIV 62
        ||: | |: || : | |: | |: | |: | |: | |: | |: | |: | |: | |: |
Db     473 YGKKADKALDRKNVTVLQGS�KH-----DGVMEVDYSNFKFYN-ASKNPNGVGVT---- 522
```

QY	63	RHSYLQE--NKIK-----TTDLHPLHDMPGNRPDVIESIDSQVLQEARPP	108
		: :	
		: :	
Db	523	NGVSHLEVGFNKVAIFNLPDLNNLAITSPVRNLDEKL-----TTKGLSPOEANKLIK	576

OY		LVSADDEIYSTKAF-----IGPIYKPPEKKRNEGRNEAHVLCINDRG	153
	:	: :	:
Dd		FUSSNKELVGKTLNFKAVADAKNTGNYDEVKKAOKDLKSURKRHEHLEKEVEKFLSKS	636

```

154 GQKEQ-----KFNSKSEI-----DNLQFYKEIEELEXE-KDGPENSCKSEPP 198
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
637 GNKNWKAQAQNSOKDEIFALINKEANRDARATAVAONI.KGTCPEI.SDQI.FNVANKI.KD 696
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

QY	199	SOEQFVPFVEGHNNGLLPDEEKKDLSNKAMPSHCDYQOQNLGNEDPKYPCNQGVITPCD	258
db	697	PKDSDEFGKNGKDFSKAFETIKALGSV-----KDTG-----	730

259 TSFTRPQWQSVYFYVPGPLSLNHLNIQSFPPNP-SNITQAODDSIQNGY 317
 IIII : : II II : :
 731 -----INPMITS-----KVENIDAAIN--FEANCAVACVGGGVMGAGS-FLVC 770

```

709      YVNNCHVNWNCMTFDONNEYTDCSENRSVSHPHSGNGCSMQDRYSYNGFCEVRERCKWDHC 377
710      | : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
711      -KVENLNLNLAALM-----EFKNKGKNRDPSKVTAQASD-----LENS- 777
712      LNFENLS

```

378 MDKHNGTDFVNNQFQE-EKLNKLOKLLIL-----LRG-LPGSGKTTLSRLLGONRDG 429
 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 16

[illegible]

484 EMKPVEVA-----IGK--GYRVEFHEPETWKPDPPELSEKRNKHG 522
KRLNKAQKQKDFNNNNNNNGLA-----NEPIAYKVNKKKAGQAAAL 907

[illegible]

573 SHNRVCVTNNH 583
:
1018 SVTQNVKAVCH 1020
1019 SVTQNVKAVCH 1020
536 LSRKVELLAQKLDLNLQAVSEAKAGFVGNLEQITDKLDRSTKHNPNNLWVESAKKVPASLS 1017

DD 1018 AKLDNYATNSH 1028
RESULT 7
00 00 00 00 00 00

US-09-925-300-968
Sequence 968, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:

APPLICANT: Craig Rosen,
APPLICANT: Steve Ruben
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101

CURRENT APPLICATION NUMBER: US/09/925,300
 CURRENT FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: PCT/US00/05988
 PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270
 PRIOR FILING DATE: 1999-03-12
 NUMBER OF SEQ ID NOS: 1890
 SOFTWARE: PatentIn Ver. 2.0

```

; SEQ ID NO 968
; LENGTH: 291
; TYPE: PRT
; ORGANISM: H
US-09-925-300-9

```

Query Match	3.4%	Score 108.5;	DB 10;	Length 291;
Best Local Similarity	23.9%	Pred. No. 0.49;		
Matches 57;	Conservative 24;	Mismatches 78;	Indels 79;	Gaps 12;

Qy 87 PGNRPVIESDSVLQEARPLVSADDEIYSTSKAFIGPIYKP-----PEKKRNEGRN 141
| | | | : | | : | | : | : |
Db 96 PQRPGVIRAL-----GPP-----GVRRRCQISPEEEERRVR 132

```
Qy      142 EAHVLNGINDRGQKQKQP-NSEKSEIDNELFQFYKEETELEKXGDGE-----NSC 193
       | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      133 ERNKAALKAKNRKRKEUFDLQAETDKLEDEKSGLRFEITELQKQRLELVIAHRPIC 192
```

QY 194 KESEPQEQFVPEYEGHNGLLKPDDEKKDLSNKAMPSCHDYQQNLGNEDPKYPC---N 249
| | : : | | : : |
Db 193 KIPEGAKGDT---GSTSTSSP-----PAC-----RPVPCISLSP 226

QY 250 GQVI-----PTFCDT-SFTSRPFWQSYPPIVPGPLPSLVHLNIQRSGPPN 299
||| || | | | | | | | | | | | |
Db 227 GPVLEALHTPLMTTPSLTPPTPSLVTP-----STPECAASHRRKSSSSGDDPS 279

RESULT 8
US-09-820-843A-107

; Sequence 107, Application US/09820843A
 ; Publication No. US20030039963A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Council of Scientific and Industrial Research

```

/ TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF
/
/ TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES
/
/ FILE REFERENCE: Q63915
/
/ CURRENT APPLICATION NUMBER: US/09/820,843A
/

```

```

; CURRENT FILING DATE: 2001-03-30
;
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 107

```

```

; LENGTH: 665
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:

```

```
; NAME/KEY: misc_feature
; OTHER INFORMATION: hypothetical protein
; NAME/KEY: misc_feature
; OTHER INFORMATION: gl|3845248
```

US-09-820-843A-107

Query Match	3.4%	Score 108;	DB 9;	Length 665;
Best Local Similarity	19.7%	Pred. No. 1.7;		

Matches	68;	Conservative	55;	Mismatches	117;	Indels	106;	Gaps	
QY	2	SYCEI-EGKFLGPREVTSPECKCLKSTES	-----	YVFHNHNSNADFHR	-----	45			

Db 28 SYNDIVEDINIKKEESSEPPFPIKKNKIDNVHTKNNQYNLNHNKSNKTHLYTGHTS 87
QY 46 -IQEKTGNDWVPVTIDV-----RGHSVLQE-NKIKTTDLHR----- 80

Db	88	FLQNTINDCVDVDNKDSINNTITKEDNNNNNNKQIEEKNKINKSDLHRQNELNQ	147
Qy	81	-----PLHDEMPGNRPDIESDSOVLQEARPPLYSADDEIYSTSKAFIG	125

```

Db      148 GKNEQDINKNEKGQDISNAENKKDYKGV-KLELEKKVEEKISDDHKKVEENKKS---203
QY      126 PTYKPPPEKKKRNE-----GRNEAHVLNGI-----NDRGGOKEQKQFNSEKSEIDNE 171

```

204 DOHKVEENKSDHKVEENKSDDHKIEVKKVEEHEDESEDKKEKSENKNKDNKDE 263

172 LFQFYKEI-----EELEKEKDGFGNSCKES-----EPSQEQFV-FFYEGHNGNL 214

[illegible]

Db 264 NDENDNDESDDEVDDEEDKNDNDIDDDKKTDTKTHLEENNEIIEKFSDDKKRNGK 323
QY 215 LK-----PDEEKKDLNSKAMPSHCDYQOONLGNEDPK 245.
Db 324 NKDTKEKSKDTEKEKSDIEKEKSKDKEKSKDKEKSKDKEK 369

RESULT 9

US-09-976-165-19
; Sequence 19, Application US/09976165
; Patent No. US20020107383A1
; GENERAL INFORMATION:
; APPLICANT: FUJIMURA, TSUTOMU
; APPLICANT: MATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; SKELETAL MUSCLE-SPECIFIC UBIOUITIN-CONUGATING ENZYME
; FILE REFERENCE: O-53599
; CURRENT APPLICATION NUMBER: US/09/976,165
; PRIOR APPLICATION NUMBER: 09/565,538
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 09/055,699
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: 08/820,170
; PRIOR FILING DATE: 1997-03-19
; PRIOR APPLICATION NUMBER: JP 63410/1996
; PRIOR FILING DATE: 1996-03-19
; PRIOR APPLICATION NUMBER: JP 69163/1997
; PRIOR FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 19
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-165-19

Query Match 3.3%; Score 106; DB 10; Length 506;
Best Local Similarity 19.9%; Pred. No. 1.7;
Matches 68; Conservative 51; Mismatches 135; Indels 88; Gaps 14;

QY 74 KTEPLHRLHDEMGNRPDVI-----ESIDSOVLQEARPLVSADDEIYSTKAFIPIYK 129
Db 135 KYALNLPFLYD---RRQITNAEYEPTEECENNSDEEFSDEVDONTPESEPL- 188
QY 130 PREKKRNEGNEAHLVINGINDRGOKKOKFNSEKSEIDNELFOFYKEIELEKEDGF 189
Db 189 -----EGEEDENPKENPEVKADEKVEPK---EIPVKDEKEVAKIEIPEVKADEKAD 237
QY 190 ENSKESSEPSQEQ---EVPYEGHNHNGLLKPDDEK-----KDLNKKAMPSHCDYQ 237
Db 238 SKDMEATPEYKEDPEKQVAKADKDEQPKATEKARAARETHRYVEERLRUSVDLKR 297
QY 238 NLGNBPDKYPCNGVIFPCDTSFSEFPEQSYRPIVAPRPL-----PSLNLHLNIQ 292
Db 298 ARKGKPKEDBK-G-IPY-----WLIYKKNVKGKPMIQAIDPEPLKLSVYS 344
QY 293 -RFGGPPRPENIFOADDSQIONGYVNNCHV-----NMNMCMTFDDONNETTDCSE 342
Db 345 LKFSKPGQPVSYTFEFH---FLPNPYFRNEVLVTKYIIKAPRDNHDPFSGMWELIEDC- 399
QY 343 NRSSVHPSGNGCSM-----QDRYVSGFCVREER 371
Db 400 -----KGCKIDRRRGKDVTVTTQTSRTATGEIIOQR 432

RESULT 10
US-09-801-368-392
; Sequence 392, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert

APPLICANT: Call, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801,368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: Patentln version 3.0
SEQ ID NO 392
LENGTH: 1093
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-801-368-392

Query Match 3.3%; Score 105; DB 10; Length 1093;
Best Local Similarity 17.3%; Pred. No. 5.8;
Matches 97; Conservative 79; Mismatches 227; Indels 158; Gaps 18;

QY 46 IOEKTGNMVPVPT--IIVRGHSYLDENKIKITDRLHPLHDEMPG----- 88
Db 60 VMRTKDMINITQVFKIAQFSKTRKRLIKESMDQHEVQGGYGFQGTWPIPLDSAK 119
QY 89 ---NRDVIESIDSOVL-----QEARPLVSADDEIYSTKA--FIGPIYKPREKKRNE 138
Db 120 FLVNYKEIIDEVYNSILTFQDPNNPPKRSKNSILKTSKGTITSSYINKPRKNS 179
QY 139 GRNEAHLVINGINDRGOKKOKFNS-EKSEIDNELFOFYKEIELEKEDFENSKESE 197
Db 180 SSSSTATTANKKG--KKNASINQNPSPLONLVFO-----T 215
QY 198 PSQEQFVYFYGHNHNGLLKPDDEKDLNKKAMPSHCDYQOONLGNBPDKYPCNGVY---P 254
Db 216 PQCFQVNSSMIMNN-----NDNHTTMNFNNDTRHNLINNISNNSMOSTIIOQOK 265
QY 255 TFCOTSFSPREMQSYRPIVAPRPLPSLNLHLNIQRFSGPRPSPNIFOADDSQIO 314
Db 266 SIHENSFNNTSATOKPLQFF-----PIPTNLQKNV-ALNPNNDNSYSNINDVYN 319
QY 315 NGYVNNCHVNMNCTFD-----QNNETTDCSENSSVHPSGNGCSMODRYVANG 364
Db 320 SSNNNNNNNNNLILVPGPMQSQOQOHHHEYLTFNNFNHSMDSITGNSKRRKRLN- 378
QY 365 FCEVAREGKCHDKNKGSTDFVYNOQFOEKLNLQKLLILRLBPSGKTLIRLLG 424
Db 379 -----QSENOQFYNOQERIQHFFILMKORPLLMQSF----- 408
QY 425 QNRDQIVSTDDYFHQOGYRYNVNQLGDAHDMONRAKOALDQGRSPVITDNTIQAME 484
Db 409 QNPND-----NH-----NEYDSNGSNNNNTVANGSIEVPSNENDNSMN 451
QY 485 MKPYEVALIGKYRVEFHEPETWKFDEDELEKRNKHGVSRRKLIQMLDRYEQMSISIV 544
Db 452 MSSRSMPFESAG-----NPSQWKLKNNKTKDQYKQITLTI 487
QY 545 MNS-----VPSHKSQRP 560
Db 488 LSSERSDVALATLTYAP 508


```

QY      308  QDDSGIQGVYVNNCHV-----NACMTEDONNETYDCSESSVHPSGCGSM 356
          :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1209  KDIRNEERERILITNONINISQERDLOKHGFHTM-----NHLHGDVSENSQINSHHG- NR 1263

QY      357  QDRYVNSG 364
          |||::|
Db      1264  QDRGNSG 1271

RESULT 13
US-10-022-819-2
: Sequence 2, Application US/10022819
: Publication No. US20030027166A1
: GENERAL INFORMATION:
:   APPLICANT: ALLEN, Antonette C. P.
:             OLSEN, Sheri J.
:             LAWRENCE, Tammy
:             ANGELLY, Tracy S.
:             RABIN, Mark B
: TITLE OF INVENTION: CODING SEQUENCE HAPLOTYPE OF THE HUMAN
:             BRCA1 GENE
: NUMBER OF SEQUENCES: 67
: CORRESPONDENCE ADDRESS:
:   ADDRESSEE: Morgan Lewis & Bockius LLP
:   STREET: 1111 Pennsylvania Avenue
:   CITY: Washington DC
:   STATE: District of Columbia
:   COUNTRY: USA
:   ZIP: 20004
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: Patent in Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/10/022,819
:   FILING DATE: 22-Apr-2002
:   CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: 09/074,452
:   FILING DATE: 1998-05-06
: ATTORNEY/AGENT INFORMATION:
:   NAME: <Unknown>
:   REGISTRATION NUMBER: <Unknown>
:   REFERENCE/DOCKET NUMBER: 044921-5049-01-US
: TELECOMMUNICATION INFORMATION:
:   TELEPHONE: 202-739-3000
:   TELEFAX: 202-739-3001
: INFORMATION FOR SEQ ID NO: 2:
:   SEQUENCE CHARACTERISTICS:
:     LENGTH: 1863 amino acids
:     TYPE: amino acid
:     STRANDEDNESS: single
:     TOPOLOGY: linear
:     MOLECULE TYPE: protein
:     HYPOTHETICAL: NO
:     ANTI-SENSE: NO
:     FRAGMENT TYPE: N-terminal
:     ORIGINAL SOURCE:
:       ORGANISM: HOMO SAPIENS
:       STRAIN: BRCA1
:       HAPLOTYPE: OM14
:     POSITION IN GENOME:
:       CHROMOSOME/SEGMENT: 17
:       MAP POSITION: 17q21
: SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-022-819-2

Query Match      3.2%; Score 104; DB 9; Length 1863;
Best Local Similarity 18.7%; Pred. No. 15;
Matches      94; Conservative 67; Mismatches 158; Indels 184; Gaps 24;
QY      65  HSYLOENIKTTDLHRPLHDEMGNRPDVIF---SIDSQVLQF-----ARP 107

```

[illegible]

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-996A-33

Query Match 3.2%; Score 104; DB 10; Length 4636;

Best Local Similarity 23.6%; Pred. No. 51;
Matches 50; Conservative 27; Mismatches 79; Indels 56; Gaps 13;

QY 187 DGFNSCKESESQGFVPEYCHNGLKPD-----EKKDLSNKAMP--CDYQQ 237
DB 3414 DG-ENDC-----GDSDELNCDFVCLSGQFKCTKNQKCIPIVNLRCNGD 3457
QY 238 NLGNEDKYPK-NGOVIPTF--CDTSTFRPEWQSVYFVYGGPPLSLNHLNIQRF 294
DB 3458 DCGDEDEDCPENSQSPDYFOCKTKHCKISLW-----VCDEDPDCADASDEANCDKK 3511
QY 295 SGPPNPPSIFAQDDSQIQNGYVNNC-HVNNCMFTDQNNNEYTCSENRSVHPSGNG 353
DB 3512 TCGPHE---FOCKN-----NNCIPDHWR-----DSQN---DCSDNSDEENCKPQT 3551
QY 354 CSMQDRYVSGFCEVRERCKWCHMDKXNGTD 385
DB 3552 CTLKDFLCANGDC-VSSRFWCGDFDCADGSD 3582

RESULT 15

US-09-825-144-14
; Sequence 14, Application US/09825144
; Patent No. US20020037286A1
; GENERAL INFORMATION:
; APPLICANT: Matthias Krause
; APPLICANT: Antonio S. Sechi
; APPLICANT: Frank B. Gertler
; APPLICANT: Jorgen Wehland
; TITLE OF INVENTION: Methods for Altering T-Cell and Macrophage Activation
; FILE REFERENCE: M0656/7065
; CURRENT APPLICATION NUMBER: US/09/825,144
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/194,215
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 819
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-825-144-14

Query Match 3.2%; Score 103; DB 10; Length 819;

Best Local Similarity 19.5%; Pred. No. 5.6;
Matches 144; Conservative 76; Mismatches 236; Indels 284; Gaps 33;

QY 15 EVTSEPRCKKLKSTESVYFHNHNSNADFHRIOEKTG-----NDWVPTVII 60
DB 72 EKPEKEPPPLKPTGGSPRGCTQNSVSRDPEVKGVLKPVSPKPTSLTKEDSKPVVLR 131
QY 61 DV--RGHSYLOENKTKTDLRPLHDEMGNRPDVIES-----IDSQVLQEA-- 105
DB 132 PPGKNLHNLNQESDLKTP-----GPKPGPAPPVPPENELKPGFSKVAGAKSKEMPAQD 184
QY 106 -----RPPLVSADDEIYSTSK-----AFIGPIYKPPPKKRN 137
DB 185 TDSKPRPRPHTFGOKPSLSTEDSOEENTSKNVPVQKSGVOLGAKSKGAPFKPKEDPED 244
QY 138 EGRNEAH-----VLNGINDRGG-----QKEKOKFNSEKSEIDNELFQ-----FYK 177
DB 245 KDHGAPSSPFGVVLKPAASRGSGGLSKNFEEKEDKTDLAKNIFLNKLNQEPAPRPPK 304
QY 178 EIEEL-----EKEKDFGNSCKESEPQGFVPEYCHNGLL-----KPDE----- 219
DB 305 APSKLTAQTPWGQSOEKEGDKNS---ATPKQALPPL-----SVLGPPPPPKPNRPPNVD 355
QY 220 ----EKKDLSNKAMPSCDIQ-QNLGNEDKYPNGQVPTFCDTSTSFREWQSVYPF 274

DB 356 LTRFRKADSANSATKSQTPTSTSLPPPPPTHPAS-----QPPLPA 396
QY 275 IVPYGPPLPSLNHLNIQRFSGPPNPISNIFOAQDDSQIQNGYVNNCHVNNCMFTDQ 334
DB 397 SUPAHPVPVSL-----PPRNKPPDLKHPIN-----DEN 426
QY 335 NEYTDSENRSVHPSGNGCSMDRYVSG--FCEVRERCKWCHMDKXNGTDKNGTDRFVNOQF 392
DB 427 QD-----GYMHSQGTG-NLEEEQESGETYEDIDSSKERDKKREKEKKLELERKE 477
QY 393 QBEKLNKLOKLLILLRGLPG-----SGKTTLSRILLGQNRD-GIVFSTD- 435
DB 478 OKERERKEQELKKFK-LTGPIQVIHAKACCDVKGKNELS-----FKQGEDIEIIRITDN 533
QY 436 -----DYFHHQDGYRYNNVNLG-----DAHDWNNQNRK 463
DB 534 PEGKWLGRTAGSYGIKTTAVEIDY-----DSLKRKKNLSLNAVPPRLVEDDDQDYYDDVAE 589
QY 464 QAI-----DQGRSPVIIDNTNIOAWENKPYVEVA 492
DB 590 QDAPNSHGSGSGGMFPPTDDEIYDGIIEEDDDGSGVPOVDEKTNANSWG- ---LKWL 646
QY 493 ICKGYRVEFHEPETWKKFPDELEKRNKHGYSRKKTAQMLDRYEQMSISIVMNSVEPSH 552
DB 647 KGKDDRKKSIREK-----PKVSESDNNEGSLPSOHKQOLDVGEE-----VYDDVDASD 694
QY 553 KSTQRPDPQGRQWRGSGSLG 572
DB 695 F-----PPPPAEMSQMSVG 709

Search completed: March 10, 2003, 18:04:49
Job time : 27 secs

Tue Mar 11 10:11:14 2003

us-09-924-654-4.ra1

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 17:48:59 ; Search time 20 seconds
(without alignments)
857.678 Million cell updates/sec

Title: US-09-924-654-4

Perfect score: 3209

Sequence: 1 MSYGELEGKFLGPREVTSE.....RQRMGSGLSGSHNRVCVTNNH 583

Scoring table: BLOSUM62

Searched: Gap0 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3209	100.0	583	2	US-08-616-392C-4
2	449.5	14.0	132	2	US-08-616-392C-10
3	372	11.6	104	2	US-08-616-392C-12
4	127	4.0	394	4	US-09-195-188-1
5	117	3.6	769	3	US-08-725-459B-38
6	112	3.5	769	3	US-08-725-459B-40
7	110	3.4	769	3	US-08-725-459B-39
8	109.5	3.4	1147	3	US-08-470-360-5
9	109.5	3.4	1147	3	US-08-471-491-5
10	109.5	3.4	1147	4	US-08-466-662-5
11	109.5	3.4	1865	1	US-08-588-985-2
12	109.5	3.4	1865	1	US-08-971-988-2
13	109.5	3.4	3289	2	US-08-477-451-2
14	109	3.4	769	3	US-08-725-459B-37
15	108.5	3.4	2391	2	US-08-446-855A-2
16	108.5	3.4	2391	2	US-09-150-741-2
17	108	3.4	769	3	US-08-725-459B-35
18	107.5	3.3	859	1	US-08-053-614-2
19	107.5	3.3	859	1	US-08-316-397B-2
20	107.5	3.3	859	1	US-08-034-306-2
21	107.5	3.3	859	4	US-09-259-437-2
22	107.5	3.3	859	5	PCT-US93-09782-2
23	107.5	3.3	1181	1	US-08-053-614-4
24	107.5	3.3	1181	1	US-08-316-397B-4
25	107.5	3.3	1181	2	US-09-034-306-4
26	107.5	3.3	1181	4	US-09-259-437-4
27	107.5	3.3	1181	5	PCT-US93-09782-4

28	107.5	3.3	1939	4	US-09-310-187A-1	Sequence 1, Appl
29	107	3.3	630	3	US-08-725-459B-17	Sequence 17, Appl
30	107	3.3	680	3	US-08-725-459B-3	Sequence 3, Appl
31	107	3.3	730	3	US-08-725-459B-2	Sequence 2, Appl
32	107	3.3	733	3	US-08-725-459B-29	Sequence 29, Appl
33	107	3.3	733	3	US-08-725-459B-30	Sequence 30, Appl
34	107	3.3	773	1	US-08-524-757-6	Sequence 6, Appl
35	107	3.3	773	3	US-08-725-459B-1	Sequence 1, Appl
36	107	3.3	773	3	US-08-725-459B-79	Sequence 79, Appl
37	107	3.3	1507	3	US-08-929-329-5	Sequence 5, Appl
38	106.5	3.3	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
39	106	3.3	506	2	US-08-820-170A-19	Sequence 19, Appl
40	106	3.3	506	3	US-09-055-699-19	Sequence 19, Appl
41	106	3.3	506	4	US-09-273-565-19	Sequence 19, Appl
42	106	3.3	506	4	US-09-565-538-19	Sequence 19, Appl
43	106	3.3	506	4	US-09-661-468-19	Sequence 19, Appl
44	105	3.3	324	2	US-08-816-755-2	Sequence 2, Appl
45	105	3.3	324	4	US-09-090-673-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-616-392C-4
Sequence 4, Application US/08616392C
Patent No. 5998165
GENERAL INFORMATION:
APPLICANT: Goold, Richard D.
APPLICANT: Akerblom, Ingrid E.
APPLICANT: Selhamer, Jeffrey
APPLICANT: Coleman, Roger
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES PANCIA
TITLE OF INVENTION: AND PANCIA ASSOCIATED WITH PANCREATIC CANCER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/616,392C
FILING DATE: 15-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/581,240
FILING DATE: 29-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0052-1US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: NMT
CLONE: 496071
US-08-616-392C-4
Query Match 100.0%; Score 3209; DB 2; Length 583;
Best Local Similarity 100.0%; Pred. No. 1e-284;

```
Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSYGEIEGFLGPREVTSPTSEPRCKKLLKSTTESYVFHNSNADFHRIOEKTGNDWVPVTII 60
DB 1 MSYGEIEGFLGPREVTSPTSEPRCKKLLKSTTESYVFHNSNADFHRIOEKTGNDWVPVTII 60
QY 61 DVGHSYLOENKIKITDLRLPHDEMGNRPDVISSQVLOEARPPPLVSADDEIYSTS 120
DB 61 DVGHSYLOENKIKITDLRLPHDEMGNRPDVISSQVLOEARPPPLVSADDEIYSTS 120
QY 121 KAFIGIYKPKPKKRNKNEGRAHVLINGIDRGQKEKQKFNSEKSEIDNELFQFYKETE 180
DB 121 KAFIGIYKPKPKKRNKNEGRAHVLINGIDRGQKEKQKFNSEKSEIDNELFQFYKETE 180
QY 181 ELEKEKDFENCKESEPQOFVFPYEGHNGNGLKPDDEKDLNKNKAMPSHCDYQOINLG 240
DB 181 ELEKEKDFENCKESEPQOFVFPYEGHNGNGLKPDDEKDLNKNKAMPSHCDYQOINLG 240
QY 241 NEPDKYPCNGQVITPFCDTSTSFPEMQSVYFIVPYGPPPLPSLYHLNIOREFSGPPNP 300
DB 241 NEPDKYPCNGQVITPFCDTSTSFPEMQSVYFIVPYGPPPLPSLYHLNIOREFSGPPNP 300
QY 301 PSNIFOAQDSDIQNGYYVNNCHVNNMNCMTFDQNNYTDCESENSSVHPGSGCQMORY 360
DB 301 PSNIFOAQDSDIQNGYYVNNCHVNNMNCMTFDQNNYTDCESENSSVHPGSGCQMORY 360
QY 361 VSGFCEVRERCKWCHMDKHNGTDRFVNOQFQEBKLNKQLKLLILLRGLPGSGKTTLSR 420
DB 361 VSGFCEVRERCKWCHMDKHNGTDRFVNOQFQEBKLNKQLKLLILLRGLPGSGKTTLSR 420
QY 421 ILLGNRQDGIIVSTDFYFHQDGYRYNNVNLGDADHWNQONRAKQAIIDOGSRVVIDNTNI 480
DB 421 ILLGNRQDGIIVSTDFYFHQDGYRYNNVNLGDADHWNQONRAKQAIIDOGSRVVIDNTNI 480
QY 481 QAWEMKPYVEVAIGRYRVEFHEPETWKFDPPELEKRNKHGVSRRKKAQMLDRYEQWS 540
DB 481 QAWEMKPYVEVAIGRYRVEFHEPETWKFDPPELEKRNKHGVSRRKKAQMLDRYEQWS 540
QY 541 ISIVNSVEPSHKSTQRPDPPOGROGWSGLSGSHNRVCVTNNH 583
DB 541 ISIVNSVEPSHKSTQRPDPPOGROGWSGLSGSHNRVCVTNNH 583

RESULT 2
US-08-616-392C-10
; Sequence 10, Application US/08616392C
; Patent No. 5998165
; GENERAL INFORMATION:
; APPLICANT: Goold, Richard D.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Seilhamer, Jeffrey
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES PANCL1A
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/616.392C
; FILING DATE: 15-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/581,240
; FILING DATE: 29-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0052-1US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0052-1US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 533948
US-08-616-392C-10
```

```
Query Match 14.0%; Score 449.5; DB 2: Length 132;
Best Local Similarity 82.1%; Pred. No. 1.7e-33;
Matches 87; Conservative 0; Mismatches 0; Indels 19; Gaps 1;
QY 1 MSYGEIEGFLGPREVTSPTSEPRCKKLLKSTTESYVFHNSNADFHRIOEKTGNDWVPVTII 60
DB 46 MSYGEIEGFLGPRE-----HHNSNADFHRIOEKTGNDWVPVTII 86
QY 61 DVGHSYLOENKIKITDLRLPHDEMGNRPDVISSQVLOEAR 106
DB 87 DVGHSYLOENKIKITDLRLPHDEMGNRPDVISSQVLOEAR 132
```

```
RESULT 3
US-08-616-392C-12
; Sequence 12, Application US/08616392C
; Patent No. 5998165
; GENERAL INFORMATION:
; APPLICANT: Goold, Richard D.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Seilhamer, Jeffrey
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES PANCL1A
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/616.392C
; FILING DATE: 15-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/581,240
; FILING DATE: 29-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0052-1US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

TOPOLOGY: linear
US-08-616-392C-12

Query Match 11.6%; Score 372; DB 2; Length 104;
Best Local Similarity 94.6%; Pred. No. 1,4e-26;
Matches 70; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSYEIGKPLGPREETVSEPRCKLKSTSTSYVFHNSNADFHRIQKGNQVPTI 60
DB 31 MSYEIGKPLGPREETVSEPRCKLKSTSTSYVFHNSNADFHRIQKGNQVPTI 90
QY 61 DVRGHSLYQENKIR 74
DB 91 DVRGHSLYQENKIR 104

RESULT 4

US-09-195-188-1
Sequence 1 Application US/09195188
Patent No. 6365722
GENERAL INFORMATION:
APPLICANT: Marcu, Kenneth B.
TITLE OF INVENTION: Y2H14 AN IKK BINDING PROTEIN
FILE REFERENCE: 178-264
CURRENT APPLICATION NUMBER: US/09/195.188
CURRENT FILING DATE: 1998-11-17
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 394
TYPE: PRT
ORGANISM: Homo sapiens
US-09-195-188-1

Query Match 4.0%; Score 127; DB 4; Length 394;
Best Local Similarity 25.5%; Pred. No. 0.0028;
Matches 41; Conservative 32; Mismatches 58; Indels 30; Gaps 7;

QY 401 QKLLILGLPGSGKTLTSLILGQNRD-----GIVSTDDYF-----HHOD----- 442
DB 81 ERIVYIMGLPGSGKTHAKLRDKEVEFGAPRVLSLDYFITEVEKEEDPDGSKRV 140
QY 443 -----GYRYNNQGLDADHNMNRA-KAIDQGRSP-VIIDNTNQAMEMKPYEVAIGK 495
DB 141 KKVREYEREA-EMETRTSMFKTKTLDGFFPFIILDAINDRVHHPDQFMSAAKTK 199
QY 496 GYRVEFHETWPKPDPELEKRNKGYSRKKIAQMLDREY 536
DB 200 GREVYLAEMSA-----DNQTCGRNIRHGKRLKELINKMADHWE 236

RESULT 5

US-08-725-459B-38
Sequence 38 Application US/08725459B
Patent No. 6084068
GENERAL INFORMATION:
APPLICANT: CONAWAY, RONALD C.
APPLICANT: CONAWAY, JOAN W.
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: SIDLEY & AUSTIN
STREET: 717 N. HARWOOD, SUITE 3400
CITY: DALLAS
STATE: TX
COUNTRY: US
ZIP: 75201-6507
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/725,459B

FILING DATE: 04-OCT-1996

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: HANSEN, EUGENIA S.

REGISTRATION NUMBER: 31,966

REFERENCE/DOCKET NUMBER: 11146/07501

TELECOMMUNICATION INFORMATION:

TELEPHONE: 214-981-3300

TELEFAX: 214-981-3400

INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:

LENGTH: 769 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: Protein

LOCATION: 1..769

OTHER INFORMATION: /note= "amino acid 1-556 and

US-08-725-459B-38

Query Match 3.6%; Score 117; DB 3; Length 769;
Best Local Similarity 19.3%; Pred. No. 0.063;
Matches 79; Conservative 48; Mismatches 160; Indels 122; Gaps 16;

QY 9 KFLGPREETVSEPRCKLKSTSTSYVFHNSNADFHRIQKTNVAVPTIIDVRGHSY 67
DB 307 KKLSPALDVASDNHFKFKHNDSEKIKSDKNKQSDVDSGRGDDPLPRAKDKVNNLK 366
QY 68 LOENKIKTTDLHRLHDEMGNRPDVIESIDQVLOEARPLVADDEIYTSKAFIOP 127
DB 367 AOEGKVFRTNS-----DRKSPGSLPKVEE-----MDMDEFEQPTMSFESYL 407
QY 128 -YKPEPKRKNENRHAIVLNGINDRGOKKOKFNSKSEIDNELFOFYIEIELEK 186
DB 408 SYDPRKKKKVYVITSGAL-----GEGLKKDKSKTS-----KNLSAOK-- 449
QY 187 DGFENCKESEPSOEQVPEYEGHNNGLKPDDEKKDLNKRAMPSCDYQONLGNPDY 246
DB 450 -----LPRANENKSD---KLOPAGAB-----PTRPRKV 474
QY 247 PCNGOVIPTFCDTSTSPREMOYPRIVYGPPLPSLNHLNQ-----RFGSPNPPS 302
DB 475 PTD-VLPALDIPLPALQTYR-----PLPSLELISSFOPRKAFFSPQEE 521
QY 303 NIFQADDSQJONGYVYVNCVNMNMCMTFDQNNETDSCENRSSVHPG----- 351
DB 522 AGFTGRMNSKMQYISGSKCALPRKMTLHQ-----QCIRNIDSLFEYGVPSVLEPYL 576
QY 352 NGCSMDRY---VSNGFCEVRERCMDKDKHNGTDGFVNOQFOEER 396
DB 577 ERCTPDQLYRIEHCNHVLIETFDLMKVHC-----HDPKREER 614

RESULT 6

US-08-725-459B-40
Sequence 40 Application US/08725459B
Patent No. 6084068
GENERAL INFORMATION:
APPLICANT: CONAWAY, RONALD C.
APPLICANT: CONAWAY, JOAN W.
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: SIDLEY & AUSTIN
STREET: 717 N. HARWOOD, SUITE 3400
CITY: DALLAS
STATE: TX
COUNTRY: US
ZIP: 75201-6507

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/725,459B
FILING DATE: 04-OCT-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: HANSEN, EUGENIA S.
REGISTRATION NUMBER: 31,966
REFERENCE/DOCKET NUMBER: 11146/07501
TELEPHONE: 214-981-3300
TELEFAX: 214-981-3400
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..769
OTHER INFORMATION: /note= "amino acids 1-564 and
OTHER INFORMATION: 569-773 of Elongin A"
US-08-725-459B-40

Query Match 3.5%; Score 112; DB 3; Length 769;
Best Local Similarity 19.3%; Pred. No. 0.18;
Matches 78; Conservative 48; Mismatches 166; Indels 112; Gaps 15;

QY 9 KFLGPREVTSRCKKLLKSTTESYVFHNHNSADFHRIQKGTG-NDWVPVTIIDVRGHSY 67
DB 307 KKLSPALDVASDNHFKPKHKHSEKIKSKDKNQSVDSVDSGRGTGDPPLRAKDKVPNNLK 366
QY 68 LQNKIKITDLHRLHDEMPGNRPDVIESIDSOVLQEARPLVSADEIYSTSKAFIGPI 127
DB 367 AQEGKVRTNS-----DRKSPGSLPKVEE-----MDMDDEFEQPTMSFESYL 407
QY 128 -YKPEKKRNEGRNEAHVNLNGINDRGQKQKFNSEKSEIDNLFQYKEIEELEKEK 186
DB 408 SYDQPRKKKKVYKTSGTAL-----GEKGLKKDKSKSTS-----KNLNSAQK-- 449
QY 187 DGFENCKESEPSEQOFPVFFYEGHNHNGLLKPDDEKKDLNKAHPSHCDYQOQNLGNEPKY 246
DB 450 -----LPKANENKSD---KLOPAGAE-----PTRPRKV 474
QY 247 PCNGQVPTFCDTSTSFPRPQSVYFIVPYGPPPLPSLNYHLNIQ-----RFSGPPNPPS 302
DB 475 PTD--VLPALPDIPALQTNVR-----PLPSLELISFQPKRKAFSSPOEEEE 521
QY 303 NIFQADDSQIQNGYYVNNCHVNNCMFTDQ-----NNEYTDCSENSSVHPNGSCSM 356
DB 522 AGFTGRMNSKMQVYSGSKCAYLPKMMTLHQOCIRVLKNIDSGGVYPSVLEPVLCTP 581
QY 357 QDRY-----VSNQCFCEVRERCKHCKMDKHNGTDRFVNOQFOEEK 396
DB 582 DQLYRIEENHVLIEETDOLMKVHC-----HRDFKEER 614

RESULT 7
US-08-725-459B-39
; Sequence 39, Application US/08725459B
; Patent No. 6084068
; GENERAL INFORMATION:
; APPLICANT: CONAWAY, RONALD C.
; APPLICANT: CONAWAY, JOAN W.
; TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:

ADDRESSEE: SIDLEY & AUSTIN
STREET: 717 N. HARWOOD, SUITE 3400
CITY: DALLAS
STATE: TX
COUNTRY: US
ZIP: 75201-6507
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/725,459B
FILING DATE: 04-OCT-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: HANSEN, EUGENIA S.
REGISTRATION NUMBER: 31,966
REFERENCE/DOCKET NUMBER: 11146/07501
TELEPHONE: 214-981-3300
TELEFAX: 214-981-3400
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..769
OTHER INFORMATION: /note= "amino acids 1-560 and
OTHER INFORMATION: 565-773 of Elongin A"
US-08-725-459B-39

Query Match 3.4%; Score 110; DB 3; Length 769;
Best Local Similarity 19.3%; Pred. No. 0.28;
Matches 78; Conservative 49; Mismatches 165; Indels 112; Gaps 15;

QY 9 KFLGPREVTSRCKKLLKSTTESYVFHNHNSADFHRIQKGTG-NDWVPVTIIDVRGHSY 67
DB 307 KKLSPALDVASDNHFKPKHKHSEKIKSKDKNQSVDSVDSGRGTGDPPLRAKDKVPNNLK 366
QY 68 LQNKIKITDLHRLHDEMPGNRPDVIESIDSOVLQEARPLVSADEIYSTSKAFIGPI 127
DB 367 AQEGKVRTNS-----DRKSPGSLPKVEE-----MDMDDEFEQPTMSFESYL 407
QY 128 -YKPEKKRNEGRNEAHVNLNGINDRGQKQKFNSEKSEIDNLFQYKEIEELEKEK 186
DB 408 SYDQPRKKKKVYKTSGTAL-----GEKGLKKDKSKSTS-----KNLNSAQK-- 449
QY 187 DGFENCKESEPSEQOFPVFFYEGHNHNGLLKPDDEKKDLNKAHPSHCDYQOQNLGNEPKY 246
DB 450 -----LPKANENKSD---KLOPAGAE-----PTRPRKV 474
QY 247 PCNGQVPTFCDTSTSFPRPQSVYFIVPYGPPPLPSLNYHLNIQ-----RFSGPPNPPS 302
DB 475 PTD--VLPALPDIPALQTNVR-----PLPSLELISFQPKRKAFSSPOEEEE 521
QY 303 NIFQADDSQIQNGYYVNNCHVNNCMFTDQ-----NNEYTDCSENSSVHPNGSCSM 356
DB 522 AGFTGRMNSKMQVYSGSKCAYLPKMMTLHQOCIRVLKNIDSGGVYPSVLEPVLCTP 581
QY 357 QDRY-----VSNQCFCEVRERCKHCKMDKHNGTDRFVNOQFOEEK 396
DB 582 DQLYRIEENHVLIEETDOLMKVHC-----HRDFKEER 614

RESULT 8
US-08-725-459B-5
; Sequence 5, Application US/08470260
; Patent No. 6077706

GENERAL INFORMATION:

APPLICANT: Covacci, Antonello
 APPLICANT: Bugnoli, Massimo
 APPLICANT: Telford, John
 APPLICANT: Macchia, Giovanni
 APPLICANT: Rappuoli, Rino
 TITLE OF INVENTION: Helicobacter pylori Proteins Useful
 TITLE OF INVENTION: for Vaccines and Diagnostics
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Chiron Corporation
 STREET: 4560 Horton Street
 CITY: Emeryville
 STATE: California
 COUNTRY: USA
 ZIP: 94608-2916
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/470,260
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/256,848
 FILING DATE: 21-OCT-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: McClung, Barbara G.
 REGISTRATION NUMBER: 33,113
 REFERENCE/DOCKET NUMBER: 0316, 001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (510) 601-2708
 TELEFAX: (510) 655-3542
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1147 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-470-260-5

Query Match 3.4%; Score 109.5; DB 3; Length 1147;
 Best Local Similarity 19.1%; Pred. No. 0.57;
 Matches 128; Conservative 98; Mismatches 240; Indels 205; Gaps 33;

QY 3 YGEIGKFLGPREVYSEPRCKKLTSTESYVFHNHNSADHRIOEKTGNDMPVTIIV 62
 DB 473 YGKKADKALDREKNVTLQSGSLKH-----DGVAFVDSNFKYTN-ASKNPNKGVGT----- 522
 QY 63 RGHSTLQD-NKIK-----TTDLHRLPHDEMGRNPVIESIDSOVLQEARP 108
 DB 523 NGVSHLEVGFKVAIFNLPLDNLNLAITSFVRNLEDKL-----TTKLSLQSEANKKLKD 576
 QY 109 LVSADDEIYSTSKAF-----IGPIYKPEPKKRNNGRNEAHVINGINDRG 153
 DB 577 FLSSNKELVGKTLNFKNAVADAKNTGNYDEVKAKQDKLESLRKREHLEKEVEKLESKS 636
 QY 154 GQKEKQ-----KFNSEKSEI-----DNELFOFYKEIELEKE-KDGFENSKSESEP 198
 DB 637 GKNKMEAKAQAQSKDEIFALINKKANRDAARAIAVAQNLGKIRELSDKLENVKNKLKD 696
 QY 199 SOEOFPVEYEGHNNGLLPDEEKDLSNKAHPSHCDYOQNLGNPDXYPCNGVVIPTFCD 258
 DB 697 FDKSFDEKRNKNDKFAETLALKGSV-----KDLG----- 730
 QY 259 TSFTSFREMOGVYPIFYGPPLPSLVNHLNIOFSGPPNP-SNIFOADDSDQIQNGY 317
 DB 731 -----INDEWIS-----KVENLNAALN--EFKNGKKNKDFSKYQAKSD--LENS- 770
 QY 318 YVNNCHVAMNCTFPDQNNETDSCENRBSVHSGNGSCMODRYVYNGCEVNERCWMKDH 377

DB 771 -VKDVII-----NQKVTDKVDNLNQ-----AVSVAKATGDFSRVEQA-----L 807
 QY 378 MDKHNGTDFRVNQFQE-EKLNKIQKLLI-----LRG-LPGSGKTTLSRILLGVRDG 429
 DB 808 ADLKNFSKEQLAQQAQKNESINARKSEIYOSKYNVNGTLVGLSQAETLTLKSNFSD 867
 QY 430 IVFSTDDYFHHODGYRRNVN-QLGDADHWNQNAKQAIIDGRSPVILIDNTINQIA-----W 483
 DB 868 I-----KELNAKLGSENNNNNGLK-----NEPIYAKVKKKAGQASL 907
 QY 484 EMKPYEVA-----IGK--GYRFEFEPETWKFDEPELEKRNKG 522
 DB 908 EEPYAGVAKKVNKIDRLNQLASGLVGQAGAFPLKR-----DKYVDLSKVG 957
 QY 523 VSR-KKIAQMLDREYQMS-----ISTVANSVEPSKSTORPP-----PGQRMGSLG 572
 DB 958 LSRNQLAOKIDNLNQAVSEAKQFQGNLBQITDKLSDTKHNPMLNMLVESAKKVPASLS 1017
 QY 573 SHNVCVTNH 583
 DB 1018 AKLDNVATNSH 1028

RESULT 9

US-08-471-491-5
 Sequence 5, Application US/08471491B
 Patent No. 6090611
 GENERAL INFORMATION:
 APPLICANT: Covacci, Antonello
 APPLICANT: Bugnoli, Massimo
 APPLICANT: Telford, John
 APPLICANT: Macchia, Giovanni
 APPLICANT: Rappuoli, Rino
 TITLE OF INVENTION: Helicobacter pylori Proteins Useful for Vaccines And
 TITLE OF INVENTION: Diagnostics
 FILE REFERENCE: CHIR0044
 CURRENT APPLICATION NUMBER: US/08/471,491B
 CURRENT FILING DATE: 1995-06-06
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 5
 LENGTH: 1147
 TYPE: PRT
 ORGANISM: Helicobacter pylori
 US-08-471-491-5

Query Match 3.4%; Score 109.5; DB 3; Length 1147;
 Best Local Similarity 19.1%; Pred. No. 0.57;
 Matches 128; Conservative 98; Mismatches 240; Indels 205; Gaps 33;

QY 3 YGEIGKFLGPREVYSEPRCKKLTSTESYVFHNHNSADHRIOEKTGNDMPVTIIV 62
 DB 473 YGKKADKALDREKNVTLQSGSLKH-----DGVAFVDSNFKYTN-ASKNPNKGVGT----- 522
 QY 63 RGHSTLQD-NKIK-----TTDLHRLPHDEMGRNPVIESIDSOVLQEARP 108
 DB 523 NGVSHLEVGFKVAIFNLPLDNLNLAITSFVRNLEDKL-----TTGLSLQSEANKKLKD 576
 QY 109 LVSADDEIYSTSKAF-----IGPIYKPEPKKRNNGRNEAHVINGINDRG 153
 DB 577 FLSSNKELVGKTLNFKNAVADAKNTGNYDEVKAKQDKLESLRKREHLEKEVEKLESKS 636
 QY 154 GQKEKQ-----KFNSEKSEI-----DNELFOFYKEIELEKE-KDGFENSKSESEP 198
 DB 637 GKNKMEAKAQAQSKDEIFALINKKANRDAARAIAVAQNLGKIRELSDKLENVKNKLKD 696
 QY 199 SOEOFPVEYEGHNNGLLPDEEKDLSNKAHPSHCDYOQNLGNPDXYPCNGVVIPTFCD 258
 DB 697 FDKSFDEKRNKNDKFAETLALKGSV-----KDLG----- 730
 QY 259 TSFTSFREMOGVYPIFYGPPLPSLVNHLNIOFSGPPNP-SNIFOADDSDQIQNGY 317

```
Db 731 -----INPEWIS-----KVENLNAALN---EFKNGKNDKDFSKVTOAKSD--LENS- 770
QY 318 YVNNCHVNNMCTFDONNEYTDCSENRSVHPGSGCSDMDRYVSNCFCEVRERCKWCHC 377
Db 771 -VKDVII-----NOKVTDKVDNLQ-----AVSVAKATGDFSRVEQA-----L 807
QY 378 MDKHNGTDRFVNOQFOE-EKLNKLOKLLIL-----LRG-LPGSGKTTLSRILLGONRDG 429
Db 808 ADLNKFSKQLAQQAOKNESLNARKKSEIYQSVKNGVNGTLVNGLSQAEATLTKNFSD 867
QY 430 IVESTDDYFHODGYRYNVN-QLGDAHDWNQNRQAQIDQGRSPVITDNTNQA-----W 483
Db 868 I-----KKELNAKLGNNNNNGLK-----NEPIYAKVKKKAGQAASL 907
QY 484 EMKPYVEVA-----IGK--GYRVEFHEPETWKKFDPPELEKRNKKG 522
Db 908 EEPITAQVAKVNAKIDRLNQIASGLGVVQGAAGFPLKRH-----DKVDDLKSYG 957
QY 523 VSR-KKIAOMLDRYEQMS-----ISIVNSVEPSHKSTQRP--POGRORWGSGLG 572
Db 958 LSRNQELAQKIDNLQAVSEAKAGFGNLEQITDKLKDSTKHNPMLNWLWESAKKVPASLS 1017
QY 573 SHNRVCVTNNH 583
Db 1018 AKLDNYATNSH 1028

RESULT 10
US-08-466-662-5
; Sequence 5, Application US/08466662B
; Patent No. 6130059
; GENERAL INFORMATION:
; APPLICANT: Covacci, Antonello
; APPLICANT: Bugnoli, Massimo
; APPLICANT: Telford, John
; APPLICANT: Macchia, Giovanni
; APPLICANT: Rappuoli, Rino
; TITLE OF INVENTION: Helicobacter Pylori Proteins Useful For Vaccines And
; FILE REFERENCE: CHIR0057
; CURRENT APPLICATION NUMBER: US/08/466,662B
; CURRENT FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1147
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-08-466-662-5

Query Match 3.4%; Score 109.5; DB 4; Length 1147;
Best Local Similarity 19.1%; Pred. No. 0.57;
Matches 128; Conservative 98; Mismatches 240; Indels 205; Gaps 33;

QY 3 YGEIEGKFLGPREEVTSRCKLKSTTESYVFNHNSADFHRIQKTDNDWVPVTIIDV 62
Db 473 YGKKADKALDREKNVTLOGSLK-----DGVMFVDYSNFKYTN-ASKNPKGVGV- 522
QY 63 RGHSYLOE--NKIK-----TTDLHRPLHDEMGNRPDVTESIDSVLOEARPP 108
Db 523 NGVSHLEVGFNKVAFNLPDLNLAITSFVRNLEDKL-----TTKGLSPQEAANKLKD 576
QY 109 LVSADDEIYSTKAF-----IGPIYKPEKKRNEGRNEAHVLNGINDRG 153
Db 577 FLSSNKELVGLTNFNKAVADAKNTGNYDEVKKAQDKLESLRKREHLEKEVEKKLESK 636
QY 154 GQKEQ-----KNSEKSEI-----DNELFOFYKEIELEKE-KDGFNSCKSESP 198
Db 637 GKNKNEAKAQAANSQKDEIFALINKEANDRAIAIYAQNLKGIKLSKLENVKNNLKD 696
QY 199 SOEQVPFYEGHNGLLKDPDEEKDLSNKAAMPCHDYOQNLGNPKYPCNGOVPTTCD 258
Db 697 FDKSDFEFGKNGKDFSKAETLTKALKGSV-----KDLG----- 730
```

```
QY 259 TSFTSRPQSWQVYPIVPGPLPSLNHLNQRSGPPNPP-SNIFOAQDSDQIQNGY 317
Db 731 -----INPEWIS-----KVENLNAALN---EFKNGKNDKDFSKVTOAKSD--LENS- 770
QY 318 YVNNCHVNNMCTFDONNEYTDCSENRSVHPGSGCSDMDRYVSNCFCEVRERCKWCHC 377
Db 771 -VKDVII-----NOKVTDKVDNLQ-----AVSVAKATGDFSRVEQA-----L 807
QY 378 MDKHNGTDRFVNOQFOE-EKLNKLOKLLIL-----LRG-LPGSGKTTLSRILLGONRDG 429
Db 808 ADLNKFSKQLAQQAOKNESLNARKKSEIYQSVKNGVNGTLVNGLSQAEATLTKNFSD 867
QY 430 IVESTDDYFHODGYRYNVN-QLGDAHDWNQNRQAQIDQGRSPVITDNTNQA-----W 483
Db 868 I-----KKELNAKLGNNNNNGLK-----NEPIYAKVKKKAGQAASL 907
QY 484 EMKPYVEVA-----IGK--GYRVEFHEPETWKKFDPPELEKRNKKG 522
Db 908 EEPITAQVAKVNAKIDRLNQIASGLGVVQGAAGFPLKRH-----DKVDDLKSYG 957
QY 523 VSR-KKIAOMLDRYEQMS-----ISIVNSVEPSHKSTQRP--POGRORWGSGLG 572
Db 958 LSRNQELAQKIDNLQAVSEAKAGFGNLEQITDKLKDSTKHNPMLNWLWESAKKVPASLS 1017
QY 573 SHNRVCVTNNH 583
Db 1018 AKLDNYATNSH 1028

RESULT 11
US-08-588-985-2
; Sequence 2, Application US/08588985
; Patent No. 5777094
; GENERAL INFORMATION:
; APPLICANT: Michiyuki MATSUDA et al.
; TITLE OF INVENTION: cDNA OF DOCK180 GENE AND DOCK180 PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,985
; FILING DATE: January 19, 1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1865 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL:
```

ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE: spleen cell of homo sapiens
US-08-588-985-2

Query Match 3.4%; Score 109.5; DB 1; Length 1865;
Best Local Similarity 18.7%; Pred. No. 1.2;
Matches 139; Conservative 97; Mismatches 277; Indels 229; Gaps 35;

```

QY 4 GEIEGKFL---GPREVTSEPRCKKLTSTESYVFNH-----HSNADF-----43
D 1074 GQHKIKFIPENVGPILEMTLIPETELRKATIP--IFEDMOCSEHSTSFQFENEITK 1131
QY 44 --HRIQETGDMVPVITIDV-----RGHSYLO---ENKIKTT-----DLHRPLDE 85
D 1132 LDHEVEGGRGEOYKVLFDKILLEHCRKHKYLAETGETFVAVYLMRLDYRTIMDE 1191
QY 86 MGRNPDIYESIDSOVLOEARRPLVSADDEIYSTKAFIGPIYKPEKKRNEGRNE---142
D 1192 NKENMSTGVVNLNFKYEIER-----EMV-----IRLYKCDLHKEDNYTEAY 1238
QY 143 ----AHYLANGIND-----RGQOKEK-----QKFNSEK-----165
D 1239 TLLHAKLKSEDEVCAVHNLQROGYQATTQGLKEQLYQELINHFQKGMMEALGK 1298
QY 166 ---SEIDNELFOFYKEIELEKENDGFENSCSESPSOEOPVPEYEGH-----NGLL--215
D 1299 ELAQYENEMFDYEQLSLELKKQAOFYENIVKIRKPDYFVAVGYGCGPFTFLGKVFI 1358
QY 216 ---KPEDEKDLKSNKAMPSHCDYOO-----NLGNEPDKYPCNGQVITPFCSTSTS 263
D 1359 YRGKEYEPRDEFEARLLQFPNAEKMTTSPRGDDIKNSP-----GOYIOCF-----1406
QY 264 FRPEWQSVYPIVYRGPPLSLNHYLN-IORF-----SGPPNPSPNIFQADDSOION 315
D 1407 VKPKLDLPKFKHRPVSQIVSF-YRVNEVOREFYSRPIRKGEKNPDNEFAMMIERTIYT 1465
QY 316 GYVYNCHVNW-----NCM-TFQDNEEYTDSCSENNSSVHPS--GNGCSM-356
D 1466 TAYKLPGLMFEVKSVMFEISPLENAIETMQLNDKINSVVOQHLDDPSLPINPLML 1525
QY 357 ----QDRVSNQGFCEVNRCKDKHDMKNGTDPRVNOQFOEKKLNKQKLL-----ILL 407
D 1526 LNGIVDPVWVGAFNYEAFETDRLQENPRA-----HEKTEKLDLAWOIPFLA 1576
QY 408 RGLPQSG-KTILSRILLQONRDGIYFSTDYFNHODGYRYVNVNOLDADHNMQNRKAOAI 466
D 1577 EGIIRHGKVTLEALRPFHERMEACFKQLEKYEKYGVRIMPSSLDDBRG---SRPRGMV 1633
QY 467 DQGRSP-----VIIDNTNIQANEMKRYVEVAILGKGYRVEFHEPETWV---KEDPE 513
D 1634 RSFTPPSSSRPLSVASVSLSDSTSPSRGSD---GFALEPLPLKMKHSHQDKLKD 1688
QY 514 ELEKRNKHSVGRKKLAQMLDREYQM---STSI-----VNV-546
D 1689 DLEKKKKKKKKKRNSKIQOIEFEKEKFTPTDISLOOSEAVIILSETISPLRQRPKQVAMV 1748
QY 547 -----SVPSHKSTQRRPPP 561
D 1749 IGSERRFVSPPSSPSQOTPPP 1770

```

RESULT 12
US-08-971-988-2
Sequence 2, Application US/08971988
Patent No. 5786461
GENERAL INFORMATION:
APPLICANT: MICHIOYUKI MATSUDA et al.
TITLE OF INVENTION: CDNA OF DOCK180 GENE AND DOCK180 PROTEIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Menderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700

CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/971.988
FILING DATE: 17-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/588.985
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1865 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE: spleen cell of homo sapiens
US-08-971-988-2

Query Match 3.4%; Score 109.5; DB 1; Length 1865;
Best Local Similarity 18.7%; Pred. No. 1.2;
Matches 139; Conservative 97; Mismatches 277; Indels 229; Gaps 35;

```

QY 4 GEIEGKFL---GPREVTSEPRCKKLTSTESYVFNH-----HSNADF-----43
D 1074 GQHKIKFIPENVGPILEMTLIPETELRKATIP--IFEDMOCSEHSTSFQFENEITK 1131
QY 44 --HRIQETGDMVPVITIDV-----RGHSYLO---ENKIKTT-----DLHRPLDE 85
D 1132 LDHEVEGGRGEOYKVLFDKILLEHCRKHKYLAETGETFVAVYLMRLDYRTIMDE 1191
QY 86 MGRNPDIYESIDSOVLOEARRPLVSADDEIYSTKAFIGPIYKPEKKRNEGRNE---142
D 1192 NKENMSTGVVNLNFKYEIER-----EMV-----IRLYKCDLHKEDNYTEAY 1238
QY 143 ----AHYLANGIND-----RGQOKEK-----QKFNSEK-----165
D 1239 TLLHAKLKSEDEVCAVHNLQROGYQATTQGLKEQLYQELINHFQKGMMEALGK 1298
QY 166 ---SEIDNELFOFYKEIELEKENDGFENSCSESPSOEOPVPEYEGH-----NGLL--215
D 1299 ELAQYENEMFDYEQLSLELKKQAOFYENIVKIRKPDYFVAVGYGCGPFTFLGKVFI 1358
QY 216 ---KPEDEKDLKSNKAMPSHCDYOO-----NLGNEPDKYPCNGQVITPFCSTSTS 263
D 1359 YRGKEYEPRDEFEARLLQFPNAEKMTTSPRGDDIKNSP-----GOYIOCF-----1406
QY 264 FRPEWQSVYPIVYRGPPLSLNHYLN-IORF-----SGPPNPSPNIFQADDSOION 315
D 1407 VKPKLDLPKFKHRPVSQIVSF-YRVNEVOREFYSRPIRKGEKNPDNEFAMMIERTIYT 1465
QY 316 GYVYNCHVNW-----NCM-TFQDNEEYTDSCSENNSSVHPS--GNGCSM-356
D 1466 TAYKLPGLMFEVKSVMFEISPLENAIETMQLNDKINSVVOQHLDDPSLPINPLML 1525

```

Qy 357 ---QDRVYVNCFCVRECHKDHCMKNGTDRFVNOQFOEKLKLOKLL-----ILL 407
Db 1526 LNGIVDPAMVGGFANYERAFKTDYRLOEHPA-----HEKIEKLKOLIAWOIPELA 1576
Qy 408 RGLPGSG-KTTLRILLQNGRDGIVFSDDDYFHHQDGYRYNNVNOIGDAHDNWNQRAKQAI 466
Db 1577 EGIRHGDGVTEALRPPHERMEACFKQKKEVEYGVIMPSSLDLDRG---SRPRSMV 1633
Qy 467 DQGRSP-----VIIDNTNIQWEMKPYVEVAIGKGYRVEFHEPFW-----KEDPE 513
Db 1634 RSFTMPSSRPLSVASVSSLDSTPSRPGSD-----GFALEPLIPKKMHSRQDKLKD 1688
Qy 514 ELEKRNHGVSRKKAQMLDYEYQM-----SISI-----VNN- 546
Db 1689 DLEKXDKKKEKRNKSHQEIFEKFRPTDISLOOSEAVILSETISPLRPORPKSQVMNV 1748
Qy 547 -----SVEPSHKSQTQPPPP 561
Db 1749 IGSERRFVSPPSSSQQTTPP 1770
RESULT 13
US-08-477-451-2
; Sequence 2, Application US/08477451
; Patent No. 5928865
; GENERAL INFORMATION:
; APPLICANT: Covacc1, Antonello
; TITLE OF INVENTION: Helicobacter Pylori Cagi Region
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,451
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0335.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2708
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3289 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-477-451-2
Query Match 3.4%; Score 109.5; DB 2; Length 3289;
Best Local Similarity 19.1%; Pred. No. 2.9;
Matches 128; Conservative 98; Mismatches 240; Indels 205; Gaps 33;
Qy 3 YGETEGFLGPREVTSEPRCKKLKSTESVFHNNHNADEPHRIQKGTGNDWVPTIIDV 62
Db 2004 YGKADRALDREKNVTLQGLKH-----DGMVFVDYNSFYTN-ASKNPNKGVGT---- 2053
Qy 63 RGHSLQOE--NKIK-----TTDLRPLHDEMPGNRPDVFIESIDSVLQEARPP 108
Db 2054 NGVSHLEVGNKVAIFNLPDLNLAITSFVRNLEDKL-----TTKGLSPQEAANKLIKD 2107

Qy 109 LVSADDEIYTSKAF-----TGPYIKPPEKKRNEGRNEAHVLNGINDRG 153
Db 2108 FLSSNKELVGTTLFNLKAVADAKNTGNYDEVKQAKQDLEKSLRKRREHLEKEVEKLESKS 2167
Qy 154 GOKKQKQ-----KFNEKSEI-----DNELQFYKEIEELEKE-KDGFENSCKESEP 198
Db 2168 GNKKMEAKAQAOKQDEIFALINKENRDRARATAYAGNLKGIKRELSDKLENNVKNLKD 2227
Qy 199 SQEQVFYFEGHNGLLKPDDEKDLNKNKAMPCHCDYQONLGNBPDKYPCNGQVPTFCD 258
Db 2228 FDKSDFEFKNGKNDFSKAEETLKALGVS-----KDLG----- 2261
Qy 259 TSFTSFPBWSQVVPFVPGPLPSLNYHNIORFSGPPNPP-SNIPQAODDSOIQNGY 317
Db 2262 -----INPEWIS-----KVENLNAALN--EFKNGKNKDFSKVTOAKSD--LENS- 2301
Qy 318 YVNNCHVNNWCMTFDQNNEXYDCSENSSVHPHSGNCSMODRYVNSGCEVRECRWDHC 377
Db 2302 -VKDVII-----NOKVTDKVDNLNQ-----AVSAKATGDFSRVEQA-----L 2338
Qy 378 MDKHNGTDRFVNOQFOE-EKLKLOKLLIL-----LRG-LPGSGKTTLSRILLQNGRDG 429
Db 2339 ADLANFSKEQLAQQAOKNESLNARKKSEIYQSVKNGVNGTLVGNLSQAEATTLKKNFSD 2398
Qy 430 IVFSTDDYFHHQDGYRYNNV-OLGDAHDNWNQRAKQAIQDGRSPVIDNTNIQA-----W 483
Db 2399 I-----KKELNAKLGNNNNNGLK-----NEPIYAKVNNKKAGQAASL 2438
Qy 484 EMKYVEVA-----IGK--CYRVEFHEPETWKFDPPELEKRNKHG 522
Db 2439 EEPYQAVAKVNNKIDRLNOLASGLGVVGOAGFPLKRR-----DKVDDLKSVG 2488
Qy 523 VSR-KKIAQMLDREYQMS-----ISIVMNSVPSHKSQTQPPPP---PQGRQWGSGLG 572
Db 2489 LSRNQLAQKIDNLNQAVSEAKAGFFGNLEQITDKLKDSTKHNPMLNWLAVESAKKVPASLS 2548
Qy 573 SHNRVCVTNNH 583
Db 2549 AKLDNYATNSH 2559
RESULT 14
US-08-725-459B-37
; Sequence 37, Application US/08725459B
; Patent No. 6084068
; GENERAL INFORMATION:
; APPLICANT: CONAWAY, RONALD C.
; APPLICANT: CONAWAY, JOAN W.
; TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SIDLEY & AUSTIN
; STREET: 717 N. HARWOOD, SUITE 3400
; CITY: DALLAS
; STATE: TX
; COUNTRY: US
; ZIP: 75201-6507
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/725,459B
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: HANSEN, EUGENIA S.
; REGISTRATION NUMBER: 31,966
; REFERENCE/DOCKET NUMBER: 11146/07501
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-981-3300
; TELEFAX: 214-981-3400

INFORMATION FOR SEQ ID NO: 37:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 769 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: Protein
 LOCATION: 1..769
 OTHER INFORMATION: /note="amino acids 1-552 and
 OTHER INFORMATION: 557-773 of Elongin A"
 US-08-725-459B-37

Query Match 3.4%; Score 109; DB 3; Length 769;
 Best Local Similarity 19.6%; Pred. No. 0.34;
 Matches 79; Conservative 48; Mismatches 165; Indels 112; Gaps 16;

Db 9 KFLGPREVTSERPCKKLKSTTESYVFHNHSMADFHRIQETG-NDWVPTIIDRGHSY 67
 307 KKLSPALDVASDNHFKKFKHDESEKIKSDKKOSYDVSRCGTGDP LPRAKDKVPPNLIK 366
 QY 68 LOENKIKITDLHRLHDEMPGNRPDIYESIDSOYLQEARPLVSADDEIYSTSKAFICPI 127
 Db 367 AQDEKVRINS-----DRKSPSLPKVEE-----MDMDDEFEGPTMSFESYL 407
 QY 128 -YRPEKKRNEGRNEAHVNLGINDRGQKQKQKFNSEKSEIDELFOFYIEIELEKEK 186
 Db 408 SYDOPRKKKKVVTSGTAL-----GEKGLKKRDKSKTS-----KNLSAOK-- 449
 QY 187 DGFENSCSESPSOQFPFYEFGNNGLLKPDDEKKDLSENKAMPSHCDYQNLGNEPKY 246
 Db 450 -----LPRANKNSD---KLQPRGAE-----PRPKKV 474
 QY 247 PCNGVPTFCDSFTSEPRMOSYPRIVYGPPLPSLNYHLNQ-----RSGGPNPSS 302
 Db 475 PTD-VLPALDIPILAQTNYR-----PLPSLELISSTQPRKAFSSQDEEE 521
 QY 303 NIFQAQDSQIQNGYYNCHVNMCMTFDQ--NNEYIDCE--NRSVHPSGNGCSM 356
 Db 522 AGFGRRMNSKMOYVSGSGCAYLPKMTLHQVLLKNNIDISIFEVGSVPYSLPEVLERCTP 581
 QY 357 QDRY-----VSNGFCEVRERCKDHGMKHNDRFVNOQFOEEK 396
 Db 582 DQLYRIECCNHVLIETDQMKVHC-----HRDFKEER 614

RESULT 15
 US-08-446-855A-2
 Sequence 2, Application US/08446855A
 Patent No. 5849573
 GENERAL INFORMATION:
 APPLICANT: Stewart, Thomas S
 APPLICANT: Flores, Maria V
 TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Nixon & Vanderhye PC
 STREET: 1100 No. 5849573th Glade Road, 8th Floor
 CITY: Arlington
 STATE: Virginia
 COUNTRY: USA
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.24
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/446,855A
 FILING DATE: 06-Jul-1995

CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Mitchard, Leonard C
 REGISTRATION NUMBER: 29,009
 REFERENCE/DOCKET NUMBER: 47-80
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-816-4000
 TELEFAX: 703-816-4100
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2391 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-446-855A-2

Query Match 3.4%; Score 108.5; DB 2; Length 2391;
 Best Local Similarity 17.2%; Pred. No. 2.2;
 Matches 105; Conservative 80; Mismatches 190; Indels 235; Gaps 24;

QY 2 SYGIEGKFLGPREVTSERPCKKLKSTTESYVFHNHSMADFHRIQETGNDWVPTIID 61
 Db 1500 TYKKNKYKDMGLNNIYDELNGSTSHSTNDHLYLDNFTSD-----EIGNN----- 1546
 QY 62 VRGSHYLOENKIKITDLH-----RPLHDEMPGNRPDIYES-----IDSOYLQ 103
 Db 1547 -----KNNDMLSEKSTISNKNPGSYIVDSYVNNKIKMKMELIDNEULN 1594
 QY 104 EAPPLVSADDEIYSTSKAFICPIYKPPKKRNEGRNEAHVNLGINDRGQKQKQKFN 163
 Db 1595 DEYNNVNMNSNYNNASAFV-----NGKDRND-----NL 1624
 QY 164 EKSEIDELFOFYIEIELEKEKGFENSC-----KESPSOQFPFYEFGH-NKLLRPD 218
 Db 1625 ENDCEKMMDHITYKHYNRLNRRSTNERMMLVNNKESNHE-----KGHRNGLNKN 1678
 QY 219 EEK-----KOLSNKAMPSH-CDYQNLGNEPKDYPCNGOVIPTFCDSFTSFRPEW 268
 Db 1679 KEKNNEKKNKDKKKNHYNNHKNNEYNNSNISKF--NNYV-----DDINKKYEDE 1732
 QY 269 QSVYPTI--VPYGPPLPSLNY-----HLNIOFSS-----GPPNPPSNIFQAQ 308
 Db 1733 NDIYFTHSQGNNDLNDNLTLSSELNTDEYDDDYDEDEDDYDDDDDDDDDD 1792
 QY 309 DDSQIQNGYYNCHVNMCMTFDQNNNEYTDCSENRSVHPSGNGCSMODRYVSGFCEV 368
 Db 1793 GEDEEDNDYNDGDYSNLSRSSRIDY-----SSVIYSGNENIFNEKYNDIGFKII 1845
 QY 369 REBCMKD-----HCMDKHNGTDRFVNOQFOEEKLKL 400
 Db 1846 DNRNEKEKEKKCFYLGCGCYRIGSSVEFDMSAIHCV-----TITRL 1889
 QY 401 OKLLILLGLPGSGKTTLSR-----ILL-----GONRGDIAF 432
 Db 1890 NKKAILINCNPETVSTVDYDESRDLKFDELTTEVIFINFENSNGVITIAFGQTSNNLVF 1949
 QY 433 STDYFHHQDGYRYVNOLGDAH-----DNNQ-----NRAQA 465
 Db 1950 SL-----YKNNVILIGSVHVKVLIYKGINFRYVILKIDQPKNNKFKLSKAIOF 2000
 QY 466 IDQGRSPVIT 475
 Db 2001 ANEVKFPVIV 2010

Search completed: March 10, 2003, 17:55:29
 Job time : 28 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 17:39:28 ; Search time 42 Seconds
(without alignments)
1849.646 Million cell updates/sec

Title: US-09-924-654-4

Perfect score: 3209
Sequence: 1 MSYGEIEGRKLGPREVTSE.....RQRMGSLGSHNVCVTNNH 583

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: A_Geneseq_101002.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	3209	100.0	583	18 AAN26654
2	3205	99.9	594	21 AAB56501
3	542	16.9	119	23 ABB22992
4	509	15.9	659	22 AAB93057
5	425.5	13.3	243	22 AAB84326
6	334	10.4	969	22 ABB63009
7	321.5	10.0	127	22 AAU28356
8	247	7.7	137	22 AAU28168
9	195.5	6.1	227	19 AAN98859
10	139.5	4.3	726	22 AAG6716

11	130.5	4.1	2485	21 AAB18172	Plasmodium falcipa
12	128	4.0	2500	17 AAB18272	Plasmodium falcipa
13	128	4.0	2913	17 AAM00384	Plasmodium falcipa
14	127	4.0	394	23 ABB76271	Human IKK binding
15	127	4.0	629	22 AAY72168	Human RNA metaboli
16	122	3.8	799	21 AAB21003	Human nucleic acid
17	121.5	3.8	1123	21 AAY58277	Heliothis armigera
18	121.5	3.8	2010	21 AAB18218	Plasmodium falcipa
19	120	3.7	1080	22 ABB11277	Novel human diagno
20	120	3.7	1111	22 AAU37362	Staphylococcus aur
21	117.5	3.6	1844	21 AAB18250	Plasmodium falcipa
22	117	3.6	702	23 ABB04333	Human zinc finger
23	117	3.6	769	19 AAB60796	Deletion mutant de
24	117	3.6	942	22 ABB62498	Drosophila melanog
25	116	3.6	1383	22 ABB63586	Drosophila melanog
26	115.5	3.6	797	22 ABB62042	Drosophila melanog
27	115.5	3.6	1247	21 AAB18215	Plasmodium falcipa
28	115	3.6	1714	21 AAB18275	Plasmodium falcipa
29	113.5	3.5	1156	22 ABB26549	Novel human diagno
30	113.5	3.5	1817	21 AAB18301	Plasmodium falcipa
31	113	3.5	417	22 AAG70780	S cerevisiae apopt
32	113	3.5	507	21 AAB18173	Plasmodium falcipa
33	113	3.5	980	21 AAB18294	Plasmodium falcipa
34	112.5	3.5	64	22 ABB16674	Novel human diagno
35	112.5	3.5	1979	21 AAB18171	Plasmodium falcipa
36	112	3.5	769	19 AAB60798	Deletion mutant de
37	111.5	3.5	662	22 ABB59430	Drosophila melanog
38	111.5	3.5	1064	22 AAM41117	Human polyptide
39	111.5	3.5	2295	21 AAB18180	Plasmodium falcipa
40	111	3.5	904	21 AAB18320	Plasmodium falcipa
41	111	3.5	1166	21 AAB18268	Plasmodium falcipa
42	111	3.5	1308	21 AAB18167	Plasmodium falcipa
43	110.5	3.4	939	22 ABB58143	Drosophila melanog
44	110.5	3.4	1373	22 ABB62368	Drosophila melanog
45	110	3.4	552	22 ABB59326	Drosophila melanog

ALIGNMENTS

RESULT 1	AAW26654	AAW26654 standard; Protein: 583 AA.
ID	AAW26654	
XX	AAW26654	
AC	16-FEB-1998	(first entry)
XX		
DE	Human PANC1A associated with pancreatic cancer.	
XX		
KW	PANC1A; pancreas cancer; diagnosis; therapy; human.	
XX		
OS	Homo sapiens.	
XX		
PN	WO9724437-A1.	
XX		
PD	10-JUL-1997.	
XX		
PF	18-DEC-1996;	96W0-US19981.
XX		
PR	15-MAR-1996;	96US-0616392.
XX		
PR	29-DEC-1995;	95US-0581240.
XX		
PA	(INCY -) INCYTE PHARM INC.	
XX		
PI	Aberblom IE, Coleman R, Goold RD, Seilhamer JJ;	
XX		
DR	WPI: 1997-363673/33.	
XX		
DR	N-PSDB; AAT91010.	
XX		
PT	DNA encoding PANC1A and PANC1B associated with pancreatic cancer -	
XX	useful for diagnosis of pancreatic cancer and identifying risk of	
PT	cancer	

```
XX PS Disclosure; Page 46-47; 65pp; English.
XX CC This protein comprises human PANC1A, which is associated with
XX CC pancreatic cancer and possibly other cancers. Its sequence was
XX CC deduced from a full-length PANC1A cDNA clone (see AAT91010) derived
XX CC from human teratocarcinoma hNT-2 cells. Specific fragments of
XX CC PANC1A (see AAT91008) and PANC1B (see AAT91009) cDNA are claimed.
XX CC Also claimed are: antisense molecules; expression vectors;
XX CC transfected host cells; and purified PANC1A and PANC1B
XX CC polypeptides. Diagnostic testing for the presence of PANC1A or
XX CC PANC1B in a sample provides a means for positively correlating
XX CC abnormal levels of PANC1A or PANC1B with pancreatic cancer
XX CC (claimed). The host cells can be cultured for the recombinant
XX CC production of PANC1A or PANC1B proteins (claimed). PANC1A and
XX CC PANC1B polypeptides can be used to screen for antagonists and
XX CC inhibitors, and to raise diagnostic antibodies.
XX SQ Sequence 583 AA;
Query Match 100.0%; Score 3209; DB 18; Length 583;
Best Local Similarity 100.0%; Pred. No. 8,9e-255;
Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSYGEIEGKFLGPREEVTSERCKKLLSTTESYVFHNHNSADFHRIQKTDNDWVPVTII 60
Db 1 MSYGEIEGKFLGPREEVTSERCKKLLSTTESYVFHNHNSADFHRIQKTDNDWVPVTII 60
Qy 61 DVRGSHYLQENKIKTTDLRHLPHDEMPCNRPDVIESDSQVLQEARPLVSADDEIYSTS 120
Db 61 DVRGSHYLQENKIKTTDLRHLPHDEMPCNRPDVIESDSQVLQEARPLVSADDEIYSTS 120
Qy 121 KAFIGPIYKPEKKRNEGRNEAHLVNGINDRGQKQKFNSEKSEIDNELFQFYKEIE 180
Db 121 KAFIGPIYKPEKKRNEGRNEAHLVNGINDRGQKQKFNSEKSEIDNELFQFYKEIE 180
Qy 181 ELEKEKDGFSNCKSESPSQEQFVPFYEGHNGLLKPDEEKKDLSNKAHPSHCDYQOQNLG 240
Db 181 ELEKEKDGFSNCKSESPSQEQFVPFYEGHNGLLKPDEEKKDLSNKAHPSHCDYQOQNLG 240
Qy 241 NEPKYPCNGQVIFTFCDTSFTSRPEWQSVYPIVPGPLPSLNHLNTORESGPPNP 300
Db 241 NEPKYPCNGQVIFTFCDTSFTSRPEWQSVYPIVPGPLPSLNHLNTORESGPPNP 300
Qy 301 PSNIFOADDSDIQNGYYVNNCHVNNWNCMTFDQNNYTDGSENSSVHPSGNGCSMDQRY 360
Db 301 PSNIFOADDSDIQNGYYVNNCHVNNWNCMTFDQNNYTDGSENSSVHPSGNGCSMDQRY 360
Qy 361 VSNGFCEVRERCWKDCHCKDNGTDFRVNQFOEKLKLLKLLGLPGSGKTTLSR 420
Db 361 VSNGFCEVRERCWKDCHCKDNGTDFRVNQFOEKLKLLKLLGLPGSGKTTLSR 420
Qy 421 ILLGQNRDGIYFSTDYFHHODGYRYNNQLGDAHDWNNQRAKQIDQGRSPVIIDNTNI 480
Db 421 ILLGQNRDGIYFSTDYFHHODGYRYNNQLGDAHDWNNQRAKQIDQGRSPVIIDNTNI 480
Qy 481 QAWEMKPYVEVAIGKGYRVEFHEPETWKKFDPDELEKRNKKGVRKKIAQMLDRIYEQMS 540
Db 481 QAWEMKPYVEVAIGKGYRVEFHEPETWKKFDPDELEKRNKKGVRKKIAQMLDRIYEQMS 540
Qy 541 ISIVWNSVEPSHKSTORPPPPQGRQGGSLGSHNRVCVTNNH 583
Db 541 ISIVWNSVEPSHKSTORPPPPQGRQGGSLGSHNRVCVTNNH 583
RESULT 2
AAB56501
ID AAB56501 standard; Protein: 594 AA.
XX AC AAB56501;
XX AC
XX DT 13-MAR-2001 (first entry)
XX XX
```

```
DE Human prostate cancer antigen protein sequence SEQ ID NO:1079.
XX KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
XX KW vulnary; gastrointestinal; nephrotropic; antineoplastic; gynaecological;
XX KW antibacterial; gene therapy; neural; immune; reproductive; renal;
XX KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX KW wound; infectious disease.
OS Homo sapiens.
XX PN WO2000055174-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US05988.
XX PR 12-MAR-1999; 99US-0124270.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.
XX PI Rosen CA, Ruben SM;
XX WPT: 2000-587513/55.
XX N-PSDB: AAF15704.
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX PS Claim 11: Page 1502-1504; 2338pp; English.
XX CC AAF15566 to AAF16505 encode the human prostate cancer associated
XX CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX CC The prostate cancer antigens can have neuroprotective, cytostatic,
XX CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
XX CC nephrotropic, antineoplastic, gynaecological and antibacterial activities,
XX CC and can be used in gene therapy. The prostate cancer antigen
XX CC polynucleotides may be used for detection of prostate cancer, chromosome
XX CC identification, as chromosome markers, and for numerous other diagnostic
XX CC or research purposes. The prostate cancer antigens may be used to treat
XX CC disorders such as neural, immune, muscular, reproductive,
XX CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX CC AAB57303 represent sequences used in the exemplification of the present
XX CC invention.
XX SQ Sequence 594 AA;
Query Match 99.9%; Score 3205; DB 21; Length 594;
Best Local Similarity 99.8%; Pred. No. 1.9e-254;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MSYGEIEGKFLGPREEVTSERCKKLLSTTESYVFHNHNSADFHRIQKTDNDWVPVTII 60
Db 12 MSYGEIEGKFLGPREEVTSERCKKLLSTTESYVFHNHNSADFHRIQKTDNDWVPVTII 71
Qy 61 DVRGSHYLQENKIKTTDLRHLPHDEMPCNRPDVIESDSQVLQEARPLVSADDEIYSTS 120
Db 72 DVRGSHYLQENKIKTTDLRHLPHDEMPCNRPDVIESDSQVLQEARPLVSADDEIYSTS 131
Qy 121 KAFIGPIYKPEKKRNEGRNEAHLVNGINDRGQKQKFNSEKSEIDNELFQFYKEIE 180
Db 132 KAFIGPIYKPEKKRNEGRNEAHLVNGINDRGQKQKFNSEKSEIDNELFQFYKEIE 191
Qy 181 ELEKEKDGFSNCKSESPSQEQFVPFYEGHNGLLKPDEEKKDLSNKAHPSHCDYQOQNLG 240
Db 192 ELEKEKDGFSNCKSESPSQEQFVPFYEGHNGLLKPDEEKKDLSNKAHPSHCDYQOQNLG 251
Qy 241 NEPKYPCNGQVIFTFCDTSFTSRPEWQSVYPIVPGPLPSLNHLNTORESGPPNP 300
Db 252 NEPKYPCNGQVIFTFCDTSFTSRPEWQSVYPIVPGPLPSLNHLNTORESGPPNP 311
```

QY 301 PSNIFQADDSQIQNGYVYNNCHVNNCMFTPDNNETDCSENRRSSVHPGNGCQMDRY 360
 DB 312 PSNIFQADDSQIQNGYVYNNCHVNNCMFTPDNNETDCSENRRSSVHPGNGCQMDRY 371
 QY 361 VSNIGFEVERRCRCKDCHMDKHNCTDRFVNOQFOEKLNLKLLILRGPGSGKTTLSR 420
 DB 372 VSNIGFEVERRCRCKDCHMDKHNCTDRFVNOQFOEKLNLKLLILRGPGSGKTTLSR 431
 QY 421 ILGQNRDGIIVFSTDDYFHHQDGYRYVNOQLGDAHDMNQRAKQALIDQGRSPYIDNTNT 480
 DB 432 ILGQNRDGIIVFSTDDYFHHQDGYRYVNOQLGDAHDMNQRAKQALIDQGRSPYIDNTNT 491
 QY 481 QAEMKPYVEVALGKGYRVEFHEPETHMKRDPPELEKRNHGVSRKKIAQMLDRYEQMS 540
 DB 492 QAEMKPYVEVALGKGYRVEFHEPETHMKRDPPELEKRNHGVSRKKIAQMLDRYEQMS 551
 QY 541 ISIVMNSVEPSHKSTORPPPOGRORWGSLSGSHNRVCYVNNH 583
 DB 552 ISIVMNSVEPSHKSTORPPPOGRORWGSLSGSHNRVCYVNNH 594
 RESULT 3
 ID ABP42992 standard; Protein; 119 AA.
 AC ABP42992:
 DT 22-AUG-2002 (first entry)
 DE Human ovarian antigen HPDMO61, SEQ ID NO:4124.
 XX
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive.
 XX
 OS Homo sapiens.
 PN WO200200677-A1.
 PD 03-JAN-2002.
 PF 07-JUN-2001; 2001MO-US18569.
 XX
 PR 07-JUN-2000; 2000US-209467P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Birse CE, Rosen CA;
 DR WPI: 2002-147878/19.
 DR N-PSDB; ABO56069.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 11, SEQ ID NO 4124; 2922pp; English.
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use

CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 119 AA:
 QY 362 SNGCFEYRRCRCKDCHMDKHNCTDRFVNOQFOEKLNLKLLILRGPGSGKTTLSRI 421
 DB 15 ASGCFEYRRCRCKDCHMDKHNCTDRFVNOQFOEKLNLKLLILRGPGSGKTTLSRI 74
 QY 422 ILGQNRDGIIVFSTDDYFHHQDGYRYVNOQLGDAHDMNQNR 461
 DB 75 LIGQNRDGIIVFSTDDYFHHQDGYRYVNOQLGDAHDMNQNR 114
 RESULT 4
 ID AAB93057 standard; Protein; 659 AA.
 AC AAB93057:
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:11862.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.
 PN EP1074617-A2.
 PD 07-FEB-2001.
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INSR.
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI: 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602

CC haemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy
 CC and Crohn's disease, atopic dermatitis, Goodpasture's syndrome, gout,
 CC multiple sclerosis, osteoarthritis, osteoporosis, psoriasis, rheumatoid
 CC arthritis or ulcerative colitis and uveitis.

XX Sequence 243 AA;

Query Match 13.3%; Score 425.5; DB 22; Length 243;
 Best Local Similarity 45.7%; Pred. No. 1.2e-26;
 Matches 84; Conservative 35; Mismatches 60; Indels 5; Gaps 2;

QY 398 NKTGLLILRLGPGSGKTLRLILGQNRDGIYFSDDYFHHODG-YRNVNOLGDAHD 456
 DB 38 HSRKRLHYLRGLPGSGKTLTLAOLQHPFRALIFSTDDEFREDGAYERNPDLERAE 97
 QY 457 WNONRAKQAIIDGRSPVITINTNIOAMKPYVEVAIGKGYRVEFHEPETMKFDEPELE 516
 DB 98 WNOGRRAKARINISPIIIDNTNLHAMEMKPYAVMALENNYEVIFRPPDRMKPFNVQELA 157
 QY 517 KRKKHGVSRKKIAQMDREYQMSISIVNNSVEPS---HKSTQRPFPFGGRQWGSGLG 572
 DB 158 RRNIHGVSRKKIRMKERYEHDYTFHSLHAERPSRNRNDRNALPLSNARYMNSYTE 217
 QY 573 SHNR 576
 DB 218 FPNR 221

RESULT 6

ABB63009 standard; Protein; 969 AA.

XX ABB63009;
 XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 15819.

XX Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001MO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL07112.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure; SEQ ID NO 15819; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB1016175), expressed DNA
 CC sequences (AB101840-AB1016175) and the encoded proteins
 CC (ABB5737-ABB72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 969 AA;

Query Match 10.4%; Score 334; DB 22; Length 969;
 Best Local Similarity 31.0%; Pred. No. 3.1e-18;
 Matches 84; Conservative 37; Mismatches 92; Indels 58; Gaps 6;

QY 320 NNCVWMCMTFDQNNNEYTCSENRS-----VHPSGNCSCMDRYVSN 363
 DB 33 NNN 92
 QY 364 GFCEVRECRKDKCMKHNKSTDFVNOQFOEELNKLQKLLILRLGPGSGKTLRLIL 423
 DB 93 SIC--QRAQAGH-----KMIITMRPGSGKSTLSESL 124
 QY 424 GON-----RDGIVFSTDYFHHODGYYRNVNOLGDAHDHNONRAKQAIIDGRSPVIT 475
 DB 125 ROHLILDRHOVRD-FVLSSDYFKTRRGYVFNPTLLPAHEMNQQRKDAAGSWPITV 183
 QY 476 DNTNIOAMKPYVEVAIGKGYRVEFHEPETMKFDEPELEKRNKGVSRKKIAQMDLRY 535
 DB 184 DNTNTWVWEMQPVQFAVRHGYIELLEPNTSWCKSASKLAQKNVNVPRENIQRLERE 243
 QY 536 EYQMSISIVNNSVEPSH-----KSTQRPFP 561
 DB 244 ERTTAGELIOLMKETKYSVELPOLRNHPPLP 274

RESULT 7

AAU28356 standard; Protein; 127 AA.

XX AAU28356;

XX 18-DEC-2001 (first entry)

DE Novel human secretory protein, Seq ID No 713.

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 XX ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 XX transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 XX amyotrophic lateral sclerosis; placental disorder; thrombocytopenia;
 XX ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 XX gut protection; lung; liver fibrosis; immune deficiency; infection;
 XX severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 XX multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 XX fertility; analgesic; pain; antigen.

XX Homo sapiens.

XX WO200166689-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001MO-US04942.

XX 07-MAR-2000; 2000US-0519705.

PR 19-MAY-2000; 2000US-0574454.

PR 17-JUN-2000; 2000US-0596193.

PR 14-JUL-2000; 2000US-0616847.

PR 19-SEP-2000; 2000US-0665363.

PR 20-OCT-2000; 2000US-0693267.

XX (HYSE-) HYSEO INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 XX WPI; 2001-589934/66.
 DR N-PSDB; AAS45256.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders -
XX
PS
PS Example 2: SEQ ID No 713; 107pp; English.
XX
XX
CC The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions,
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAU28020-AAU28395 represent novel human secreted protein
CC amino acid sequences of the invention.
XX
SQ Sequence 127 AA;

Query Match 10.0%; Score 321.5; DB 22; Length 127;
Best Local Similarity 53.9%; Pred. No. 1.6e-18;
Matches 62; Conservative 18; Mismatches 34; Indels 1; Gaps 1;

Qy 398 NKLOKLLILLRGLPGSGKTTLSRILLGNRDGIVFTDTHFDHDDG-VRYNNVNLGDHDD 456

Db 12 HSFPRKHLVLLRGLPGSGKTTLSRILLGNRDGIVFTDTHFDHDDG-VRYNNVNLGDHDD 71

Qy 457 WNQNRKQAIIDQGRSPVITDNTNIQAWEMKPYVEVAIGKGYRVEFPEHETWKFED 511

Db 72 WNQNRKQAIIDQGRSPVITDNTNIQAWEMKPYVEVAIGKGYRVEFPEHETWKFED 126

RESULT 8

AAU28168

ID AAU28168 standard; Protein: 137 AA.

AC AAU28168;

XX 18-DEC-2001 (first entry)

DE Novel human secretory protein, Seq ID No 337.

XX Human, secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW gut protection; lung; liver fibrosis; immune deficiency; infection;
KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;

KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW fertility; analgesic; pain; antigen.
XX Homo sapiens.
OS
PN WO200166689-A2.
XX
XX 13-SEP-2001.
PD
XX 05-MAR-2001; 2001WO-US04942.
XX
XX 07-MAR-2000; 2000US-0519705.
XX
XX 17-JUN-2000; 2000US-0574454.
PR
XX 17-JUN-2000; 2000US-0596193.
PR
XX 14-JUL-2000; 2000US-0616847.
PR
XX 19-SEP-2000; 2000US-0665363.
PR
XX 20-OCT-2000; 2000US-0693267.
PR
XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
XX WPI: 2001-589934/66.
DR N-PSDB: AAS45068.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders -
XX

PS Example 5: SEQ ID No 337; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions,
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAU28020-AAU28395 represent novel human secreted protein
CC amino acid sequences of the invention.

XX Sequence 137 AA;

Query Match 7.7%; Score 247; DB 22; Length 137;
Best Local Similarity 42.9%; Pred. No. 2.3e-12;
Matches 48; Conservative 21; Mismatches 39; Indels 4; Gaps 1;

Qy 469 GRSPVITDNTNIQAWEMKPYVEVAIGKGYRVEFPEHETWKFEDPEELEKRNKHGVSRRKI 528

DB 4 GISPITIDNTNLMAMEKPYAVVALENNYEVIFREPDTKPFNVQELARNIHCVSRKI 63
 QY 529 AQMLDRYQMSISIVNSVEPS---HKSTORPPPOGRQSGSLGSHNR 576
 DB 64 HRMKERYEHVTFHSVLHAEKPSMRNPNRNALPSNNARYMNSYTEFPNR 115

RESULT 9

AAW98859
 ID AAW98859 standard; Protein; 227 AA.

AC AAW98859;

DT 31-MAR-1999 (first entry)

DE H. pylori GHPO 1662 protein.

KM GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
 KM peptic ulcer disease.

OS Helicobacter pylori.

PN MO9843478-A1.

PD 08-OCT-1998.

PF 01-APR-1998; 98WO-US06371.

PR 29-JUL-1997; 97US-0902615.

PR 01-APR-1997; 97US-0833457.

PR 24-JUN-1997; 97US-0881227.

PA (HUMA-) HUMAN GENOME SCI INC.

PI AL-Garawi A, Kleantous H, Miller C, Oomen RP, Tomb J;

DR WPI; 1998-542293/46.

DR N-PSDB; AAX14578.

PT New isolated Helicobacter polynucleotides - used to develop products

PT for the diagnosis, prevention and treatment of Helicobacter

PT infections and gastrointestinal diseases

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

RESULT 10
 AAG66716
 ID AAG66716 standard; Protein; 726 AA.
 AC AAG66716;

DT 02-NOV-2001 (first entry)

DE Human transcription regulatory factor 80.

KM Human; transcription regulatory factor 80; bZIP domain; RING finger;

KM zinc finger; recombinant production; malignant tumour; cancer;

KM blood disease; HIV infection; human immunodeficiency virus;

KM immune disorder; inflammatory condition; cytostatic; anti-HIV;

KM antiinflammatory; immunomodulator.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Domain 18..55

FT Domain /label= RING_finger

FT Domain /note= "C3HC4 zinc finger domain"

FT Domain 102..127

FT Domain /label= bZIP_domain_1

FT Domain 178..208

FT Domain /label= bZIP_domain_2

FT Domain 219..256

FT Domain /label= bZIP_domain_3

PN WO200155381-A1.

PD 02-AUG-2001.

PF 21-JAN-2001; 2001MO-CN00082.

PR 28-JAN-2000; 2000CN-0111582.

PA (BIOD-) BIODOR GENE TECHNOLOGY LTD SHANGHAI.

PI Mao Y, Xie Y;

DR WPI; 2001-483247/52.

DR N-PSDB; AAH76684.

PT Human transcription regulatory factor 80 and encoded polynucleotide,

PT applicable in diagnosis and treatment of malignant tumor, hemopathy,

PT HIV infection, immunological diseases and various inflammations

XX

XX

XX

XX

XX

XX

Query Match 4.3%; Score 139.5; DB 22; Length 726;
 Best Local Similarity 18.0%; Pred. No. 0.019;

Sequence 726 AA;

Query Match 6.1%; Score 195.5; DB 19; Length 227;
 Best Local Similarity 32.0%; Pred. No. 8.5e-08; Mismatches 63; Indels 25; Gaps 8;

Sequence 227 AA;

Query Match 6.1%; Score 195.5; DB 19; Length 227;
 Best Local Similarity 32.0%; Pred. No. 8.5e-08; Mismatches 63; Indels 25; Gaps 8;

	Matches	129;	Conservative	97;	Mismatches	234;	Indels	255;	Gaps	30;
Qy	14	REEVTS	PRCK--LKSTTES	VFNHNS	ADPHRIQ	EKTGN	DWPVTI	IDVRGHS	VLQENK	72
Db	56	RVPITP	PNCKE	IIGGISE	SEPM	LSTH-			VRKH-LR	91
Qy	73	IKTTDL	HRPLH	DMPGN	RNDVIE		-----	STDSQV		107
Db	92	LEL--	LHKEYE	DEIDC	LQKEVEE	LKSNLS	ESQIK	TILDPL	TLVQGN	149
Qy	108	PLVSAD		-----	DEIYST	KAFIG	PIYPPE	KKRNEG		148
Db	150	SKINP	ETVAEW	KKLR	ANETI	YEVK	DDVKK	EANKK	LKLLENG	209
Qy	149	INDRG		-----	QKEKQ	KFNSE	KSEIDN	ELFQV	KEETE	198
Db	210	SPQFGR	FAYAA	LQSKVE	OYERET	NRLK	KALERS	DKYIEE	LSQVAQ	269
Qy	199		-----	SEQFV	PFVEG	HNGLL	KDEE	KKDL	SNRAMP	237
Db	270	ICQTAL	SADG	KGSGSE	DVYSK	NOG--	DSARK	QCSST	SSSSHL	328
Qy	238		-----	NLGN	EPDKY	PCNGQ	IVPT	CTDTSF	-TSFRE	282
Db	329	ESTSKA	DLNCS	KNDL	QOEQ	VEVMD	VTDT	MDYLER	EWGNK	388
Qy	283	PSLNYH	NIORF	SGPP	NPSP	IFIAQ	DDSQ	FQNGY	VYVNN	342
Db	389	PCTPLS	LSC	LQ	LSTPEN	RESSV	QAGG	SKHSN		435
Qy	343	NRSSVH	PSGNC	S	MODRY	VSNG	FC	VEVR	ERCWK	395
Db	436	NVSNKD	SSEDD	ISRSN		-----	EKKSE	CFSP	KTFW	483
Qy	396	KLNLQK	LILL	RGLP	SGSK	TTL	SRILL	QGNR	DGIV	455
Db	484	EGNTIAN	SV		-----	GEISS	SKLS		-----	515
Qy	456	DWNONR	AKA	QAI	DG	RSP	VIID		-----	496
Db	516	SFENR	RTS	SEAS	MDAA	YLDK	ISEL	DSM	SESD	574
Qy	497	YRVF	HEP	ETW	KW	KD	PELE	KRNK	HGVS	533
Db	575	--E	FFLE		-----	EPKLE	KTEL	NLSK	SGSL	624
Qy	534	RYEQM	SIS	IVM	NSVEP		-----	SHKST	QRP	571
Db	625	DFSU	HSS	CPV	TNEI	KPP	SC	LFQ	TEFSQ	672
RESULT	11									
AAB18172										
ID	AAB18172	standard;	Protein;	2485	AA.					
XX	AAB18172;									
XX										
DT	07-NOV-2000	(first entry)								
XX										
XX										
DE	Plasmodium falciparum	chromosome 2	related protein	SEQ ID NO:29.						
XX										
KW	Plasmodium falciparum;	chromosome 2;	human malaria	parasite;	vaccine;					
KW	antimalarial;	malaria;	protozoa	infection;	insecticide.					
XX										
OS	Plasmodium falciparum.									
XX										
PN	WO2000025728-A2.									
XX										


```
PD 31-OCT-1996.
XX
PF 26-APR-1996; 96WO-US05798.
XX
PR 27-APR-1995; 95US-0430908.
XX
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Baruch DI, Howard RJ, Pasloske BL;
XX
DR WPI; 1996-497376/49.
DR N-PSDB; AAT41852.
XX
PT New Plasmodium falciparum erythrocyte membrane proteins - used to
PT develop products for the diagnosis, treatment or prevention of
PT malaria parasite infections
XX
PS Claim 1; Figure 12; 149pp; English.
XX
CC A polypeptide comprising a Plasmodium falciparum (Pf) erythrocyte
CC membrane protein 1 (PfEMP1) or active fragments or analogues of that
CC protein can be used in the treatment or prevention of symptoms of a
CC malaria parasite infection. The polypeptides can inhibit, block or
CC reverse the sequestration of erythrocytes in patients suffering from
CC malaria. Nucleic acids derived from the PfEMP1 gene can be used as
CC probes and primers to identify a Plasmodium falciparum parasite, the
CC primers used to generate characteristic amplification patterns from
CC different P. falciparum strains. Antibodies specifically
CC immunoreactive with the PfEMP1 polypeptide or its fragments may be
CC used in diagnosis of malaria infection. This is the PfEMP1 protein
CC of the MC type of Plasmodium falciparum. An alternative, truncated
CC PfEMP1 protein is given in AA00385.
XX
SQ Sequence 2913 AA;
Query Match 4.0%; Score 128; DB 17; Length 2913;
Best Local Similarity 20.8%; Pred. No. 1.3; Indels 126; Gaps 22;
Matches 88; Conservative 56; Mismatches 153;
QY 117 YTSKAFIGPIYKPPKKRNGRNEAHVLNGIN----- 150
DB 1461 YNLLKHKIDPCI-----KREKQDKTEHKCINGCNKCECVKRWLETKGNGWNIKKHYNI 1515
QY 151 DRGQKQKQKFNSEKSEINELQF-YKEIEE-LEKEKGFENSEKSESPSQEQFVPFYE 208
DB 1516 NSNDDKETIAYNKSYFVDQGLFDTDYKKAQKVVEDEK-----BRKKIWGCT 1562
QY 209 GHNNGLLKPDDEKDLNKNAMPCHCDYQNLGNEDPKYPCNGQVPTFCDTSTSPREP 268
DB 1563 GHDECEKEKENKNITNLI---SELQKITSCQKNHPNGK---TACD-PFPSTPPE- 1614
QY 269 QSYVFPFIVPGPLPSNLNHLNIQRFSGPPNPP-----SNIFQAQDSDSIQNGYVNN 321
DB 1615 -ETDPLDDDDTDPDLD-DOHTQPKFCPPPPPPPTCVEXIAKELRVEAEGKINNELKNG 1672
QY 322 CHVNNWMTFDQNNETDCSENRSSVHPSGNGSMQDRVVSNGFCFVRECRKWDHCKDH 381
DB 1673 KDFNGKCNVKNKNGAVIGES-----CKFEQTY-ENSVNNINNK-----KD 1714
QY 382 NGTDFE-VNQOFOEELNKLQKLIL-----LRGLPGSGKTTLS----- 419
DB 1715 QNERFKIGCKWNFKYIGTIRKDLCPRRHEHMCDDLMSLMTTISDALLKKIQEAA 1774
QY 420 -----RILLQNRDGIIVSTDDYFHHQ--DGVRNNYNQLGD---AHD-WNQNRKOA 465
DB 1775 KSERDDIIRKLEQN-----SCDE---HRICDAMKYSADLGDIIIRGDLNKNKSKQG 1825
QY 466 IQD 468
DB 1826 LQK 1828
RESULT 14
```

```
ABB76271
ID ABB76271 standard; Protein; 394 AA.
XX
AC ABB76271;
XX
DT 12-AUG-2002 (first entry)
XX
DE Human IKK binding protein Y2H14.
XX
Y2H14; IKK binding protein; I kappa B kinase; cytostatic;
KW antiinflammatory; immunosuppressive; immunostimulant; human.
XX
OS Homo sapiens.
XX
PN US365722-B1.
XX
PD 02-APR-2002.
XX
PF 17-NOV-1998; 98US-0195188.
XX
PR 17-NOV-1998; 98US-0195188.
XX
PA (UYNV ) UNIV NEW YORK STATE RES.
XX
PI Marcu KB;
XX
WPI; 2002-442213/47.
DR N-PSDB; ABL57342.
XX
PT Polynucleotide encoding an I kappa B kinase binding protein Y2H14 and
PT the recombinant protein encoded for elucidating and controlling
PT pathways leading to inflammation and apoptosis
XX
PS Claim 1; Column 11-14; 9pp; English.
XX
CC The present sequence is the protein sequence for a human I kappa B
CC kinase (IKK) binding protein, designated Y2H14. Y2H14 was shown,
CC using yeast two-hybrid screens, to bind to the C-terminal domain
CC regions of IKK-alpha and IKK-beta, specifically to the region of
CC the IKK proteins made up of the contiguous helix-loop-helix and
CC leucine zipper domains. The Y2H14 protein and other IKK binding
CC proteins are useful for elucidating and controlling pathways
CC leading to inflammation and apoptosis. They can also be used to
CC detect IKK complexes and modulate IKK activity in cells undergoing
CC signalling by inflammatory mediators, and to identify
CC therapeutically active agents that modulate the binding or
CC interaction of Y2H14 with IKK-alpha or -beta. Molecules that
CC prevent the formation of Y2H14/IKK complexes or inhibit the
CC dissociation of these complexes are useful for boosting the
CC immune system, or as immunosuppressives, or as antiinflammatory
CC agents.
XX
SQ Sequence 394 AA;
Query Match 4.0%; Score 127; DB 23; Length 394;
Best Local Similarity 25.5%; Pred. No. 0.082;
Matches 41; Conservative 32; Mismatches 58; Indels 30; Gaps 7;
QY 401 OKLLILLRGLPGSGKTTLSRILLGNRD-----GVFSTDDYF-----HHOD----- 442
DB 81 ERIVIMRGLPGSGKTHVAKLRDRKEVEFGPGAPRVLSLDDYFITEVEKEEDPSGKKV 140
QY 443 -----GYRYNNQLGDAHDWNNRA-KQAIDGRSP-VIIDNTNIQAWEMKPYVEAIGK 495
DB 141 KKKVEYEVEA-EMETYRTSMFKTKFLDDGFFPFILDAINDRVHRHFDQFWSAAATK 199
QY 496 GYRVEFHPETWKKDPDELEKRNKHGVSRRKIAQMLDRIE 536
DB 200 GFEVYLAEMSA----DNQTCGRNIHGRKLKEINKMADHWE 236
RESULT 15
AAI72168
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 18:04:15 ; Search time 10.733 Seconds
(without alignments)
55.005 Million cell updates/sec

Title: US-09-924-654-4_COPY_404_417

Perfect score: 69

Sequence: 1 LILLRGLPGSGKTT 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188354 seqs, 42170167 residues

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCT08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69	100.0	583	10 US-09-924-654-4	Sequence 4, Appl
2	69	100.0	594	10 US-09-925-300-1079	Sequence 1079, Ap
3	50	72.5	92	9 US-10-156-239-29	Sequence 29, Appl
4	50	72.5	92	9 US-10-156-239-30	Sequence 30, Appl
5	50	72.5	92	9 US-10-156-239-42	Sequence 42, Appl
6	50	72.5	92	10 US-09-795-693-29	Sequence 29, Appl
7	50	72.5	92	10 US-09-795-693-30	Sequence 30, Appl
8	50	72.5	92	10 US-09-795-693-42	Sequence 42, Appl
9	50	72.5	191	10 US-09-815-242-11348	Sequence 11348, A
10	50	72.5	235	10 US-09-803-286A-4	Sequence 4, Appl
11	50	72.5	479	10 US-09-971-309-64	Sequence 64, Appl
12	49	71.0	448	10 US-09-815-242-11562	Sequence 11562, A
13	47	68.1	215	10 US-09-815-242-5542	Sequence 5542, Ap
14	47	68.1	224	10 US-09-815-242-12227	Sequence 12227, A
15	47	68.1	224	10 US-09-815-242-12824	Sequence 12824, A
16	47	68.1	457	10 US-09-815-242-11948	Sequence 11948, A
17	47	68.1	462	10 US-09-815-242-10965	Sequence 10965, A
18	47	68.1	547	9 US-09-738-626-5755	Sequence 5755, Ap
19	46	66.7	65	10 US-09-756-996-6	Sequence 6, Appl

20	46	66.7	192	9 US-09-974-879-193	Sequence 193, App
21	46	66.7	453	10 US-09-815-242-10263	Sequence 10263, A
22	46	66.7	501	10 US-09-815-242-14089	Sequence 14089, A
23	46	66.7	748	10 US-09-949-192-43	Sequence 43, Appl
24	46	66.7	751	10 US-09-815-242-5832	Sequence 5832, Ap
25	46	66.7	825	10 US-09-815-242-12963	Sequence 12963, A
26	46	66.7	869	10 US-09-815-242-10623	Sequence 10623, A
27	46	66.7	891	10 US-09-949-192-45	Sequence 45, Appl
28	46	66.7	897	10 US-09-949-192-49	Sequence 49, Appl
29	45	65.2	212	10 US-09-815-242-13219	Sequence 13219, A
30	45	65.2	213	10 US-09-815-242-4889	Sequence 4889, Ap
31	45	65.2	231	10 US-09-815-242-10536	Sequence 10536, A
32	45	65.2	360	9 US-09-738-626-5332	Sequence 5332, Ap
33	45	65.2	459	9 US-09-738-626-5303	Sequence 5303, Ap
34	45	65.2	472	10 US-09-815-242-4945	Sequence 4945, Ap
35	45	65.2	477	10 US-09-815-242-10861	Sequence 10861, A
36	44	63.8	164	9 US-09-738-626-5292	Sequence 5292, Ap
37	44	63.8	191	9 US-09-895-913A-336	Sequence 336, App
38	44	63.8	213	10 US-09-815-242-10129	Sequence 10129, A
39	44	63.8	213	10 US-09-815-242-14070	Sequence 14070, A
40	44	63.8	215	10 US-09-815-242-5382	Sequence 5382, Ap
41	44	63.8	221	10 US-09-815-242-12195	Sequence 12195, A
42	44	63.8	307	9 US-10-108-605-183	Sequence 183, App
43	44	63.8	466	10 US-09-815-242-13525	Sequence 13525, A
44	44	63.8	523	10 US-09-815-242-13254	Sequence 13254, A
45	44	63.8	523	10 US-09-814-041A-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-924-654-4
; Sequence 4, Application US/09924654
; Patent No. US20020146712A1
; GENERAL INFORMATION:
; APPLICANT: Goold, Richard D.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Selhammer, Jeffrey J.
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: TUMOR SUPPRESSOR
; FILE REFERENCE: PC-0049 CIP
; CURRENT APPLICATION NUMBER: US/09/924, 654
; CURRENT FILING DATE: 2001-08-07
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020146712A1 496071CD1
US-09-924-654-4

Query Match      100.0%; Score 69; DB 10; Length 583;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LILLRGLPGSGKTT 14
Db      404 LILLRGLPGSGKTT 417

RESULT 2
US-09-925-300-1079
; Sequence 1079, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Ruben
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
```

```

RESULT 4
US-10-156-239-30
; Sequence 30, Application US/10156239
; Publication No. US20030036074A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: No. US20030036074A1el Nucleic Acid Sequences Encoding Human Tr
; TITLE OF INVENTION: Anpase Molecule, A Human Ubiquitin Hydrolase-Like Molecule, A
; TITLE OF INVENTION: Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefor
; FILE REFERENCE: 35800/247645
; CURRENT APPLICATION NUMBER: US/10/156,239
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 09/795,693
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/809,557
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/192,018
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,568
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/191,790
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,767
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/191,781
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-10-156-239-30

Query Match 72.5%; Score 50; DB 9; Length 92;
Best Local Similarity 71.4%; Pred. No. 0.65;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LILLRGLPGSGKTT 14
Db 4 VLLVGPPGSGKTT 17
::||| | |||||
;

RESULT 5
US-10-156-239-42
; Sequence 42, Application US/10156239
; Publication No. US20030036074A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: No. US20030036074A1el Nucleic Acid Sequences Encoding Human Tr
; TITLE OF INVENTION: Anpase Molecule, A Human Ubiquitin Hydrolase-Like Molecule, A
; TITLE OF INVENTION: Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefor
; FILE REFERENCE: 35800/247645
; CURRENT APPLICATION NUMBER: US/10/156,239
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 09/795,693
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/809,557
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/192,018
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,568
; PRIOR FILING DATE: 2001-03-14

```

```

; PRIOR APPLICATION NUMBER: 60/191,790
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,767
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/191,781
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 92
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-10-156-239-42

Query Match          72.5%; Score 50; DB 10; Length 92;
Best Local Similarity 71.4%; Pred. No. 0.65;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LILLRGLPGSGKTT 14
   :||| | |||||
Db 4 VVLLVGPSPGSGKTT 17

RESULT 6
US-09-795-693-29
; Sequence 29, Application US/09795693
; Patent No. US20020068710A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and
; FILE REFERENCE: 35800/209292
; CURRENT APPLICATION NUMBER: US/09/795,693
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 92
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-09-795-693-29

Query Match          72.5%; Score 50; DB 10; Length 92;
Best Local Similarity 71.4%; Pred. No. 0.65;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LILLRGLPGSGKTT 14
   :||| | |||||
Db 4 VVLLVGPSPGSGKTT 17

RESULT 7
US-09-795-693-30
; Sequence 30, Application US/09795693
; Patent No. US20020068710A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and
; FILE REFERENCE: 35800/209292
; CURRENT APPLICATION NUMBER: US/09/795,693
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
```

```

; LENGTH: 92
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-09-795-693-30

Query Match          72.5%; Score 50; DB 10; Length 92;
Best Local Similarity 71.4%; Pred. No. 0.65;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LILLRGLPGSGKTT 14
   :||| | |||||
Db 4 VVLLVGPSPGSGKTT 17

RESULT 8
US-09-795-693-42
; Sequence 42, Application US/09795693
; Patent No. US20020068710A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and
; FILE REFERENCE: 35800/209292
; CURRENT APPLICATION NUMBER: US/09/795,693
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 92
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-09-795-693-42

Query Match          72.5%; Score 50; DB 10; Length 92;
Best Local Similarity 71.4%; Pred. No. 0.65;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LILLRGLPGSGKTT 14
   :||| | |||||
Db 4 VVLLVGPSPGSGKTT 17

RESULT 9
US-09-815-242-11348
; Sequence 11348, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11348
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11348

Query Match 72.5%; Score 50; DB 10; Length 191;
Best Local Similarity 71.4%; Pred. No. 1.3;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 LILLRGLPGSGKTT 14
| | | | | | | | | |
Db 4 LFLIAGPGSGKTT 17

RESULT 10
US-09-803-286A-4
; Sequence 4, Application US/09803286A
; Patent No. US20020062504A1
; GENERAL INFORMATION:
; APPLICANT: Tanksley, Steven D.
; APPLICANT: Brommonschenkel, Sergio H.
; TITLE OF INVENTION: PLANT GENE CONFERRING RESISTANCE TO TOSPOVIRUSES
; FILE REFERENCE: 19603/3201
; CURRENT APPLICATION NUMBER: US/09/803,286A
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/188,356
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Lycopersicon var.
US-09-803-286A-4

Query Match 72.5%; Score 50; DB 10; Length 235;
Best Local Similarity 64.3%; Pred. No. 1.7;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 LILLRGLPGSGKTT 14
| | | | | | | | | |
Db 4 VISITGMPGSGKTT 17

RESULT 11
US-09-971-309-64
; Sequence 64, Application US/09971309
; Patent No. US20020106675A1
; GENERAL INFORMATION:
; APPLICANT: UEMORI, Takashi
; APPLICANT: SATO, Yoshiaki
; APPLICANT: FUJITA, Tomoko
; APPLICANT: MIYAKE, Kazuo
; APPLICANT: MURAI, Hiroyuki
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
; FILE REFERENCE: 1422-0494P
; CURRENT APPLICATION NUMBER: US/09/971,309
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 09/446,504
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: PCT/JP98/02845

; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: JP 9-187496
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: JP 9-320692
; PRIOR FILING DATE: 1997-11-27
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-09-971-309-64

Query Match 72.5%; Score 50; DB 10; Length 479;
Best Local Similarity 76.9%; Pred. No. 3.4;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 ILLRGLPGSGKTT 14
: | | | | | | | | | |
Db 46 LLLAGPGSGKTT 58

RESULT 12
US-09-815-242-11562
; Sequence 11562, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.01A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11562
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11562

Query Match 71.0%; Score 49; DB 10; Length 448;
Best Local Similarity 64.3%; Pred. No. 4.5;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 LILLRGLPGSGKTT 14
: | | | | | | | | | |
Db 96 VVLMAGLOGSGKTT 109
US-09-815-242-5542
RESULT 13


```
Sequence 5542, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5542
LENGTH: 215
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5542
```

```
Query Match      68.1% Score 47; DB 10; Length 215;
Best Local Similarity 69.2% Pred. No. 4.3;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 2 ILRLGPGSGKTT 14
    1:1||||:|1|
Db 3 IILMGLPGAGKGT 15
```

```
RESULT 14
US-09-815-242-12227
Sequence 12227, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
```

```
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12227
LENGTH: 224
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12227

Query Match      68.1% Score 47; DB 10; Length 224;
Best Local Similarity 69.2% Pred. No. 4.4;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 2 ILRLGPGSGKTT 14
    1:1||||:|1|
Db 12 IILMGLPGAGKGT 24
```

```
RESULT 15
US-09-815-242-12824
Sequence 12824, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12824
LENGTH: 224
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12824
```

```
Query Match      68.1% Score 47; DB 10; Length 224;
Best Local Similarity 69.2% Pred. No. 4.4;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 2 ILRLGPGSGKTT 14
    1:1||||:|1|
Db 12 IILMGLPGAGKGT 24
```

Search completed: March 10, 2003, 18:19:50

Job time : 11.7333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 17:54:59 ; Search time 11.6667 Seconds
(Without alignments)
35.308 Million cell updates/sec

Title: US-09-924-654-4_COPY_404_417
Perfect score: 69
Sequence: 1 LILLRGLPGSGKTT 14

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pcp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	100.0	583	2	US-08-616-392C-4
2	56	81.2	394	4	US-09-195-188-1
3	50	72.5	479	4	US-09-446-504-64
4	50	72.5	479	4	US-09-712-286-64
5	50	72.5	806	1	US-07-980-528-2
6	50	72.5	1255	3	US-08-947-823-3
7	50	72.5	1257	3	US-08-947-823-5
8	47	68.1	222	4	US-09-134-001C-3289
9	47	68.1	815	4	US-09-134-001C-3614
10	46	66.7	65	4	US-09-268-364-6
11	45	65.2	223	4	US-09-574-141A-69
12	45	65.2	229	4	US-09-134-001C-4113
13	45	65.2	318	2	US-08-671-947-2
14	45	65.2	670	2	US-08-366-547-2
15	45	65.2	670	2	US-08-366-547-2
16	44	63.8	181	1	US-08-292-858B-2
17	44	63.8	233	4	US-08-418-444A-7
18	44	63.8	233	4	US-09-161-662-2
19	44	63.8	485	4	US-08-993-825-2
20	44	63.8	523	2	US-09-134-001C-4294
21	44	63.8	523	2	US-08-923-772-2
22	44	63.8	523	4	US-09-385-287-2
23	44	63.8	652	4	US-08-987-123-2
24	43	62.3	154	4	US-09-134-001C-3873
25	43	62.3	221	4	US-09-291-170A-8
26	43	62.3	308	4	US-09-724-884-8
27	43	62.3	341	4	US-09-149-476-607
					Sequence 3650, Ap

28	43	62.3	399	4	US-09-149-476-474	Sequence 474, App
29	43	62.3	456	2	US-08-933-750C-5	Sequence 5, Appli
30	43	62.3	456	4	US-09-234-613-5	Sequence 5, Appli
31	43	62.3	456	4	US-08-268-364-21	Sequence 21, Appli
32	43	62.3	503	3	US-08-911-853-7	Sequence 7, Appli
33	43	62.3	503	4	US-09-479-409-7	Sequence 7, Appli
34	43	62.3	503	4	US-09-479-453-7	Sequence 7, Appli
35	43	62.3	806	4	US-08-999-774A-6	Sequence 6, Appli
36	43	62.3	2763	3	US-08-486-944-2	Sequence 2, Appli
37	42	60.9	28	2	US-08-724-354D-6	Sequence 6, Appli
38	42	60.9	28	3	US-09-270-984A-6	Sequence 6, Appli
39	42	60.9	239	4	US-08-858-207A-341	Sequence 341, App
40	42	60.9	264	4	US-09-035-382-4	Sequence 4, Appli
41	42	60.9	267	4	US-09-134-001C-3224	Sequence 3224, Ap
42	42	60.9	317	1	US-08-221-750A-13	Sequence 13, Appli
43	42	60.9	338	4	US-09-134-001C-4795	Sequence 4795, Ap
44	42	60.9	368	4	US-09-134-001C-5436	Sequence 5436, Ap
45	42	60.9	406	2	US-08-222-719-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-616-392C-4
Sequence 4, Application US/08616392C
Patent No. 5998165
GENERAL INFORMATION:
APPLICANT: Goold, Richard D.
APPLICANT: Akerblom, Ingrid E.
APPLICANT: Selhammer, Jeffrey
TITLE OF INVENTION: NOVEL POLYPEPTIDES PANCIA
TITLE OF INVENTION: AND PANCIA ASSOCIATED WITH PANCREATIC CANCER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/616,392C
FILING DATE: 15-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/581,240
FILING DATE: 29-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0052-10US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
STRANDEDNESS: single
IMMEDIATE SOURCE:
LIBRARY: NMT
CLONE: 496071
US-08-616-392C-4
Query Match 100.0%; Score 69; DB 2; Length 583;
Best Local Similarity 100.0%; Pred. No. 0.0022;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LILLRGLPGSGKTT 14
:|||||
Db 404 LILLRGLPGSGKTT 417

RESULT 2

US-09-195-188-1
; Sequence 1, Application US/09195188
; Patent No. 6365722
; GENERAL INFORMATION:
; APPLICANT: Marcu, Kenneth B.
; TITLE OF INVENTION: Y2H4 AN IKK BINDING PROTEIN
; FILE REFERENCE: 178-264
; CURRENT APPLICATION NUMBER: US/09/195,188
; CURRENT FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-195-188-1

Query Match 81.2%; Score 56; DB 4; Length 394;
Best Local Similarity 69.2%; Pred. No. 0.18;
Matches 9; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LILLRGLPGSGKT 13
:|||||
Db 84 VVIMRGLPGSGKT 96

RESULT 3

US-09-446-504-64
; Sequence 64, Application US/09446504
; Patent No. 6218150
; GENERAL INFORMATION:
; APPLICANT: UEMORI, Takashi
; APPLICANT: SATO, Yoshimi
; APPLICANT: FUJITA, Tomoko
; APPLICANT: MIYAKE, Kazue
; APPLICANT: MUKAI, Hiroyuki
; APPLICANT: ASADA, Kiyoza
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
; FILE REFERENCE: 1422-408PCT
; CURRENT APPLICATION NUMBER: US/09/446,504
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: PCT/JP98/02845
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: JP 9-187496
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: JP 9-320692
; PRIOR FILING DATE: 1997-11-27
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-09-446-504-64

Query Match 72.5%; Score 50; DB 4; Length 479;
Best Local Similarity 76.9%; Pred. No. 2;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ILLRGLPGSGKTT 14
:|||
Db 46 LLAGPPGSGKTT 58

RESULT 4

US-09-712-266-64
; Sequence 64, Application US/09712266
; Patent No. 6331358
; GENERAL INFORMATION:
; APPLICANT: UEMORI, Takashi
; APPLICANT: SATO, Yoshimi
; APPLICANT: FUJITA, Tomoko
; APPLICANT: MIYAKE, Kazue
; APPLICANT: MUKAI, Hiroyuki
; APPLICANT: ASADA, Kiyoza
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
; FILE REFERENCE: 1422-408PCT
; CURRENT APPLICATION NUMBER: US/09/712,266
; CURRENT FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 09/446,504
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: PCT/JP98/02845
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: JP 9-187496
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: JP 9-320692
; PRIOR FILING DATE: 1997-11-27
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-09-712-266-64

Query Match 72.5%; Score 50; DB 4; Length 479;
Best Local Similarity 76.9%; Pred. No. 2;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ILLRGLPGSGKTT 14
:|||
Db 46 LLAGPPGSGKTT 58

RESULT 5

US-07-980-528-2
; Sequence 2, Application US/07980528
; Patent No. 5457026
; GENERAL INFORMATION:
; APPLICANT: Drevfuss, Gideon
; APPLICANT: Kiledjian, Megarditch
; APPLICANT: Portman, Douglas S.
; TITLE OF INVENTION: METHODS OF PROMOTING INTERMOLECULAR
; TITLE OF INVENTION: INTERACTIONS INVOLVING A NUCLEIC ACID AND MOLECULES
; TITLE OF INVENTION: USEFUL
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: NO. 5457026ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25; mb/MD
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/980,528
; FILING DATE: 19921120
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark

REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: UPN-0847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 806 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-980-528-2

Query Match 72.5%; Score 50; DB 1; Length 806;
Best Local Similarity 57.1%; Pred No. 3.4;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 L1LRGLPGSGKTT 14
DB 480 VYMGIGLPGAGKTT 493

RESULT 6
US-08-947-823-3
Sequence 3, Application US/08947823
Patent No. 6114605
GENERAL INFORMATION:
APPLICANT: Williamson, Valerie M.
APPLICANT: Kaloshian, Isagouhi
APPLICANT: Yaghoobi, Jafar
APPLICANT: Bodeau, John
APPLICANT: Milligan, Stephen
TITLE OF INVENTION: Procedures and Materials for Confering
TITLE OF INVENTION: Pest Resistance in Plants
NUMBER OF SEQUENCES: 5
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/947,823
FILING DATE: 09-OCT-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/18802
FILING DATE: 09-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/028,191
FILING DATE: 10-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-070210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-947-823-3
Query Match 72.5%; Score 50; DB 3; Length 1255;

Best Local Similarity 64.3%; Pred. No. 5.5;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 L1LRGLPGSGKTT 14
DB 544 VISITGMPSGKTT 557

RESULT 7
US-08-947-823-5
Sequence 5, Application US/08947823
Patent No. 6114605
GENERAL INFORMATION:
APPLICANT: Williamson, Valerie M.
APPLICANT: Kaloshian, Isagouhi
APPLICANT: Yaghoobi, Jafar
APPLICANT: Bodeau, John
APPLICANT: Milligan, Stephen
TITLE OF INVENTION: Procedures and Materials for Confering
TITLE OF INVENTION: Pest Resistance in Plants
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/947,823
FILING DATE: 09-OCT-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/18802
FILING DATE: 09-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/028,191
FILING DATE: 10-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-070210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1257 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-947-823-5
Query Match 72.5%; Score 50; DB 3; Length 1257;
Best Local Similarity 64.3%; Pred. No. 5.5;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 L1LRGLPGSGKTT 14
DB 545 VISITGMPSGKTT 558

RESULT 8
US-09-134-001C-3289
Sequence 3289, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3289

; LENGTH: 222
; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3289

Query Match 68.1%; Score 47; DB 4; Length 222;

Best Local Similarity 69.2%; Pred. No. 2.6; Mismatches 2; Indels 2; Gaps 0;

Matches 9; Conservative 2; Mismatches 2; Indels 2; Gaps 0;

QY 2 LLRLGLPGSGKTT 14

Db 10 ILLMGLPGAGKGT 22

RESULT 9

US-09-134-001C-3614

; Sequence 3614, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3614

; LENGTH: 815
; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3614

Query Match 68.1%; Score 47; DB 4; Length 815;

Best Local Similarity 64.3%; Pred. No. 11; Mismatches 3; Indels 2; Gaps 0;

Matches 9; Conservative 3; Mismatches 3; Indels 2; Gaps 0;

QY 1 LLRLGLPGSGKTT 14

Db 360 IMLTGGPGTGKTT 373

RESULT 10

US-09-268-364-6

; Sequence 6, Application US/09268364A

; Patent No. 6204063

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Lee, Jian Ming

; APPLICANT: Lightner, Jonathan

; APPLICANT: Odell, Joan

; TITLE OF INVENTION: PLANT GLYCOLYSIS AND RESPIRATION ENZYMES

; FILE REFERENCE: BB-1154

; CURRENT APPLICATION NUMBER: US/09/268,364A

; CURRENT FILING DATE: 1999-03-15

; EARLIER APPLICATION NUMBER: 60/079,387

; EARLIER FILING DATE: March 16, 1998

; NUMBER OF SEQ ID NOS: 21

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 6
; LENGTH: 65

; TYPE: PRT

; ORGANISM: Oryza sativa

US-09-268-364-6

Query Match 66.7%; Score 46; DB 4; Length 65;

Best Local Similarity 75.0%; Pred. No. 1; Mismatches 1; Indels 2; Gaps 0;

Matches 9; Conservative 1; Mismatches 1; Indels 2; Gaps 0;

QY 3 LLRLGLPGSGKTT 14

Db 35 LLHGPPGTGKTT 46

RESULT 11

US-09-574-141A-69

; Sequence 69, Application US/09574141A

; Patent No. 6395490

; GENERAL INFORMATION:

; APPLICANT: Gonsalves, Dennis

; APPLICANT: Meng, Baozhong

; TITLE OF INVENTION: RUPESTRIS STEM PITTING ASSOCIATED VIRUS

; TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND THEIR USES

; FILE REFERENCE: 07678/035005

; CURRENT APPLICATION NUMBER: US/09/574,141A

; CURRENT FILING DATE: 2000-05-18

; PRIOR APPLICATION NUMBER: 60/047,147

; PRIOR FILING DATE: 1997-05-20

; PRIOR APPLICATION NUMBER: 60/069,902

; PRIOR FILING DATE: 1997-12-17

; PRIOR APPLICATION NUMBER: 09/081,320

; PRIOR FILING DATE: 1998-05-19

; NUMBER OF SEQ ID NOS: 97

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 69

; LENGTH: 223

; TYPE: PRT

; ORGANISM: Rupestris stem pitting associated virus

US-09-574-141A-69

Query Match 65.2%; Score 45; DB 4; Length 223;

Best Local Similarity 61.5%; Pred. No. 5.5; Mismatches 3; Indels 2; Gaps 0;

Matches 8; Conservative 3; Mismatches 3; Indels 2; Gaps 0;

QY 2 LLRLGLPGSGKTT 14

Db 24 IVHAPGSGKTT 36

RESULT 12

US-09-134-001C-4113

; Sequence 4113, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 4113

; LENGTH: 229

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4113

Query Match 65.2%; Score 45; DB 4; Length 229;

Best Local Similarity 69.2%; Pred. No. 5.7;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ILRLGLPGSGKTT 14
Db 42 IILNGASGSGKTT 54

RESULT 13

US-08-671-947-2
Sequence 2, Application US/08671947
Patent No. 5827515
GENERAL INFORMATION:
APPLICANT: Shahab Reyoso, Mitra
APPLICANT: Yamamoto, Takashi
APPLICANT: Cooper, Nicole H.
APPLICANT: Kaiman, Sue S.
TITLE OF INVENTION: BACILLUS THURINGIENSIS SPOKULATION GENE
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: SANDOZ AGRO, INC.
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/671,947
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Marcus Wyner, Lynn
REGISTRATION NUMBER: 34,869
REFERENCE/DOCKET NUMBER: 133-0724
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/354-3588
TELEFAX: 415/857-1125
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-671-947-2

Query Match 65.2%; Score 45; DB 2; Length 318;
Best Local Similarity 61.5%; Pred. No. 8.1;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ILRLGLPGSGKTT 14
Db 92 ILRLGNPGTGKTT 104

RESULT 14

US-08-366-547-2
Sequence 2, Application US/08366547
Patent No. 5843737
GENERAL INFORMATION:
APPLICANT: Chen, Ian Bo
APPLICANT: Bao, Shideng
TITLE OF INVENTION: A NEW CANCER ASSOCIATED GENE, PROTEIN
TITLE OF INVENTION: EXPRESSED THEREFROM AND USES THEREOF
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS
ADDRESSEE: 6 CUSHMAN
STREET: 130 Water Street

CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/366,547
FILING DATE: 30-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Eisenstein, Ronald I.
REGISTRATION NUMBER: 30628
REFERENCE/DOCKET NUMBER: 45072
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 523-3400
TELEFAX: (617) 523-6440
TELEX: 200281 STRE UR
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 670 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-366-547-2

Query Match 65.2%; Score 45; DB 2; Length 670;
Best Local Similarity 57.1%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILRLGLPGSGKTT 14
Db 121 ILRLTGPCCGKTT 134

RESULT 15
US-09-292-858B-2
Sequence 2, Application US/09292858B
Patent No. 6455681
GENERAL INFORMATION:
APPLICANT: Dean, Frank
APPLICANT: O'Donnell, Michael E.
TITLE OF INVENTION: DNA MOLECULES ENCODING SINGLE STRAND GAP RESPONSE
TITLE OF INVENTION: PROTEINS INVOLVED IN ACTIVATION OF A DNA REPAIR/CELL
FILE REFERENCE: 22221/1011
CURRENT APPLICATION NUMBER: US/09/292,858B
CURRENT FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: 60/082,020
PRIOR FILING DATE: 1998-04-16
NUMBER OF SEQ. ID NOS: 27
SOFTWARE: Patentin Ver. 2.1
SEQ. ID NO. 2
LENGTH: 670
TYPE: PRT
ORGANISM: Homo sapiens
US-09-292-858B-2

Query Match 65.2%; Score 45; DB 4; Length 670;
Best Local Similarity 57.1%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILRLGLPGSGKTT 14
Db 121 ILRLTGPCCGKTT 134

Search completed: March 10, 2003, 18:05:21
Job time : 12.6667 secs